

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу**

«На правах рукопису»
УДК 004.9:519.233

«До захисту допущено»

В.о. завідувача кафедри ММСА

О.Л Тимошук.

«___» _____ 20__ р.

Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності 122 Комп'ютерні науки

на тему: «Моделювання нелінійних нестационарних процесів методами
інтелектуального аналізу даних»

Виконав :

студент II курсу, групи КА-74мн

Пінчук Владислав Олександрович

Керівник:

професор кафедри ММСА

д.т.н., проф. Бідюк П.І.

Рецензент:

професор ФТІ НТУУ КПІ ім. Сікорського

д.т.н. О.Є.Архипов

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.
Студент _____

Київ
2019

РЕФЕРАТ

Магістерська дисертація: 130 с., 21 рис., 18 табл., 2 додатки, 24 джерела.

МЕРЕЖА БАЙЄСА, МЕТОД ГРУПОВОГО ВРАХУВАННЯ АРГУМЕНТІВ, МЕТОДИ ПОДІБНИХ ТРАЄКТОРІЙ, НЕЙРОННА МЕРЕЖА, ПРОГНОЗУВАННЯ, НЕСТАЦІОНАРНИЙ ПРОЦЕС, ЧАСОВИЙ РЯД.

Об'єкт дослідження: нестационарні процеси різної природи в економіці та фінансах.

Предмет дослідження: інтелектуальні методи прогнозування, методи математичного моделювання.

Метою даної магістерської дисертації є порівняльний аналіз, побудова математичних моделей і функцій прогнозування для фінансово-економічних процесів різної природи за допомогою інтелектуальних методів аналізу даних.

В роботі проведено огляд інтелектуальних методів побудови моделей нелінійних нестационарних процесів. Виконано аналіз проблем, пов'язаних з математичним моделюванням процесів різної природи, представлених часовими рядами; проведено критичний огляд сучасних методів моделювання та побудови прогнозів для часових рядів.

Здійснено моделювання та оцінка короткострокових прогнозів нелінійних нестационарних процесів за допомогою нейронних та байєсівських мереж, методу групового урахування аргументів, методу подібних траєкторій.

На основі проведених досліджень запропоновано СППР з використанням досліджених методів.

ABSTRACT

Master's thesis: 130 p., 21 fig., 18 tables, 2 annexes, 24 sources.

BAYESIAN NETWORK, FORECASTING, GROUP METHOD OF DATA HANDLING, NEAREST NEIGHBOUR ALGORITHM, NEURAL NETWORK , PREDICTION FUNCTION, TIME SERIES

Object of research: nonlinear nonstationary processes of different origination in economics and finances.

Subject of research: intellectual forecasting methods, mathematical modeling methods.

The purpose of this master's thesis is comparison analysis of mathematical models and forecasting functions for financial processes of different origination using intellectual data analysis methods.

The paper reviews the intellectual methods of data analysis for financial and economic processes modeling. Made problem analysis related to mathematical modeling of processes represented as time series. Described review of modern modeling and forecasting methods for time series.

Also described and demonstrated the use of selected methods for modeling and short-time forecasting of real economic processes and comparative analysis of the results.

Based on gained results suggested a decision support system using methods described in this research.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ	8
ВСТУП	9
РОЗДІЛ 1 АНАЛІЗ РОЗВИТКУ СТАТИСТИЧНО-ЙМОВІРІСНИХ МЕТОДІВ ПРОГНОЗУВАННЯ	11
1.1. Актуальність задачі моделювання	11
1.2 Актуальність задачі прогнозування	13
1.3 Короткий огляд моделей прогнозування	16
1.3.1 Оптимальний предиктор і його властивості	16
1.3.2 Регресійні моделі	17
1.3.3 Прогнозування за допомогою нейронних мереж	20
1.3.4 Лінійні параметричні моделі	22
1.4 Критерії якості моделей та оцінок прогнозів	23
1.5 Висновки до розділу та постановка задачі	29
РОЗДІЛ 2 ВИБІР ТА ОПИС МЕТОДІВ СТАТИСТИЧНО-ЙМОВІРІСНОГО ПРОГНОЗУВАННЯ	32
2.1. Мережа Байєса	32
2.2 Прогнозування за допомогою нейронних мереж	36
2.3 Метод експоненційного згладжування	42
2.3.1 Методи вибору α	44
2.3.2 Вибір початкового наближення	45
2.4 Метод групового врахування аргументів	46
2.4.1 Основні ідеї методу	46
2.4.2 Зовнішні критерії оптимальності	47
2.4.3 Опис алгоритмів селекції	49

	7
2.5 Метод подібних траєкторій	53
2.6. СППР	58
2.6.1 Актуальність створення СППР	58
2.6.2 Опис СППР	59
2.6.3 Архітектура СППР	62
2.7 Висновки до розділу 2	67
РОЗДІЛ 3 ВИКОНАННЯ І АНАЛІЗ ОБЧИСЛЮВАЛЬНИХ ЕКСПЕРИМЕНТІВ	69
3.1 Опис даних для прогнозування	69
3.2 Опис програмного продукту	74
3.3 Побудова та порівняння оцінок короткострокових прогнозів	74
3.4 Висновки до розділу 3	85
РОЗДІЛ 4 РОЗРОБКА СТАТАП-ПРОЕКТУ	88
4.1 Інформаційна карта проекту	88
4.2 Команда стартап-проекту	90
4.3 Маркетингова стратегія та маркетинговий план проекту	91
4.4 Елементи фінансової моделі	99
4.5 Стратегія конкурентної поведінки	102
4.6 Висновки до розділу 4	104
ПЕРЕЛІК ПОСИЛАНЬ	107
ДОДАТОК А Лістинг програми	110
ДОДАТОК Б Часові ряди	114

ПЕРЕЛІК СКОРОЧЕНЬ

АКФ– Автокореляційна функція

АР–Авторегресія

АРІКС– Авторегресія з інтегрованим ковзним середнім

АРКС– Авторегресія з ковзним середнім

САПП– Середня абсолютна похибка в процентах

ЕЗ– Експоненційне згладжування

ЕОМ– Електронна обчислювальна машина

КД– Коефіцієнт детермінації

КС– Ковзне середнє

МАП– Максимальна абсолютна похибка

МГВА– Метод групового врахування аргументів

МіАПП– Мінімальна абсолютна похибка в процентах

ММП– Метод максимальної правдоподібності

МНК– Метод найменших квадратів

МПТ– Метод подібних траєкторій

ННП Нелінійний нестационарний процес

РР– Різницеве рівняння

СКП– Середньоквадратична похибка

СПП– Середня похибка

СПП– Середня похибка в процентах

ФЕП– Фінансово-економічний процес

ВСТУП

Дослідження фінансово-економічних процесів (ФЕП) сьогодні є дуже актуальним, адже очевидною є залежність усіх сфер сучасного життя від економічних процесів, тому дуже важливим є усвідомлення та грамотний аналіз цих явищ. Статистично-ймовірнісні методи є одними з найпопулярніших для прогнозування розвитку економічних процесів, а також об'ємів торгових операцій, об'ємів виробництва, формування бюджетів підприємств та держави, оцінювання альтернативних економічних стратегій, прогнозування та менеджменту економічних і фінансових ризиків. Сьогодні без економічних прогнозів не обійтись також для визначення цілей розвитку суспільства, обґрунтування напрямків розвитку економіки або передбачення наслідків прийнятих рішень.

Зрозуміло, що вже існує чимало розробок у цій галузі, тож крім проблеми розробки нових якісних методів аналізу часових рядів фінансово-економічної сфери постає проблема адекватного вибору якісних методів дослідження з уже наявних. Важливим аспектом наведених проблем є складність обчислення та побудови моделей через те, що майже завжди такі процеси підвладні впливу великої кількості випадкових факторів, пов'язаних між собою невідомим чином, і через це часто є нелінійними та нестационарними. Для забезпечення високої адекватності математичних, статистичних та ймовірнісних моделей, а також високоякісних прогнозів необхідно критично аналізувати, розвивати і вдосконалювати існуючі методи.

В роботі основна увага приділяється проблемам вибору адекватних моделей для часових рядів фінансово-економічної сфери та побудови якісних короткострокових оцінок прогнозів.

Виконано аналіз проблем, пов'язаних з математичним моделюванням процесів різної природи, представлених часовими рядами; проведено

критичний огляд сучасних методів моделювання та побудови прогнозів для часових рядів.

РОЗДІЛ 1

АНАЛІЗ РОЗВИТКУ СТАТИСТИЧНО-ЙМОВІРНІСНИХ МЕТОДІВ ПРОГНОЗУВАННЯ

1.1. Актуальність задачі моделювання

З безлічі процесів у суспільстві, фінансові та економічні є одними з найбільш поширених. До них відносять інвестиційні фінансові процеси, процеси глобалізації та державної інтеграції, курси валют, акцій і облігацій, процеси виникнення міжнародних фінансових потоків, формування та використання бюджетів на державному рівні, фінансово-економічну діяльність підприємств, біржові процеси формування цін на активи різного характеру та ще багато інших. Сучасні дослідження в галузі фінансів та економіки розвиваються в тому числі через впровадження сучасних математичних методів, і не без появи новітніх комп'ютерних технологій, що зробили можливим дослідження складних явищ і процесів. Додати частину.

Проте складності додає те, що більшість процесів зазвичай мають нелінійний нестационарний характер, при будь-якому рівні розвитку економіки. Це пояснюється наявністю тісних структурних взаємозв'язків між економіками різних країн, значними коливаннями цін на енергоносії, нестабільністю курсів валют, значними коливаннями попиту на промислові товари внаслідок нестабільного рівня життя, впливом несприятливих погодних умов на сільськогосподарське виробництво та деякими іншими факторами [1, 2]. Нестационарність проявляється у формі трендів різних порядків і змінної у часі дисперсії процесу (гетероскедастичності). В сучасному світі, особливо в економічній та фінансовій сфері, точні прогнози необхідні для того, щоб здійснювати керування цілими компаніями. Тому, звісно, що той, хто використовує кращі способи для аналізу зашумлених та хаотичних часових рядів може сподіватися на більший прибуток, оскільки значно менше піддається впливу різних криз в економіці.

Одним з відомих підходів до прогнозування процесів в економічних, фінансових та багатьох інших видах систем сьогодні є прогнозування за допомогою статистичних моделей. Найбільше вони використовуються для оцінювання середньострокових прогнозів для державних, обласних бюджетів, інвестицій та доходів, прогнозування об'ємів виробництва та нагромадження продукції, оцінювання альтернативних фінансово-економічних стратегій розвитку, прогнозування та менеджменту всеможливих ризиків, а також розв'язання ще багатьох інших задач планування й управління. Прогноз завжди здійснюється з деякою похибкою, що залежить від моделі прогнозу, яка використовується, й повноти вхідних даних. При прогнозуванні, наприклад, українського споживчого ринку, на жаль, необхідно брати до уваги його нестабільність, а також особливості поведінки українських споживачів (наприклад, покупки в період інфляції).

Також при вирішенні цих задач, особлива увага приділяється побудові короткострокових прогнозів. Сьогодні існує багато різноманітних методів для короткострокового прогнозування і більшість з них базується на аналізі минулих даних, в той час, як на ринку вся доступна інформація негайно відбивається в цінах на фінансові активи. Отже, бачимо що, аналіз фінансових та економічних процесів є досить трудомістким, а число можливих методів прогнозування зазвичай дуже обмеженим.

Аналіз методів короткострокового прогнозування виявив, що найчастіше використовуються класичні методи, такі як регресійні алгоритми, методи, що базуються на експоненційному згладжуванні, нейронні мережі, методи групового врахування аргументів тощо.

Регресійні рівняння, котрі є найпоширенішим інструментом для прогнозування, дозволяють ефективно використовувати не більше декількох десятків змінних. При застосуванні більшої кількості змінних їх зашумленість та взаємозалежність роблять практично неможливою коректну параметризацію моделі, що породжує проблему пошуку альтернативних методів прогнозування.

Суттєвий вплив на економічні процеси відіграють кількісні фактори. Проте варто брати до уваги, що до важливих особливостей задач фінансового прогнозування потрібно також віднести наявність якісних факторів, що призводить до ускладнення математичної формалізації моделі. Саме тому, швидких обертів протягом останніх десятиліть набрав розвиток нового підходу до ймовірнісно-статистичного моделювання, який ґрунтується на теорії графів, теорії ймовірностей, а також на методах прикладної статистики. Моделі, утворені за допомогою об'єднання вищезгаданих методів, називають мережами Байєса. Сьогодні вони широко використовуються для моделювання в біоінформатиці (генетичні мережі, структура білків), медицині, класифікації документів, обробці зображень, обробці даних, машинному навчанні, прогнозуванні та системах підтримки прийняття рішень, особливо для фінансової та економічної сфери. Вони широко застосовуються для обробки статистичних даних, представлених часовими рядами та перерізами, а також якісних даних у формі експертних оцінок, лінгвістичних змінних та багато іншого. Окремим класом для їх використання є системи підтримки прийняття рішень, особливо для бізнесу та маркетингу, для якісного прогнозування, класифікації даних, розпізнавання тощо.

1.2 Актуальність задач прогнозування

Будь-яка фінансова чи економічна діяльність супроводжується накопиченням великої кількості бізнес-інформації, що потребує обробки та грамотного прийняття рішення. Одним із можливих шляхів є прогнозування. Прогнозом називають науково обґрунтоване судження про можливий стан об'єкта у майбутньому, про альтернативні напрямки та терміни його здійснення. Процес розроблення прогнозів називають прогнозуванням [3]. Фінансово-економічне прогнозування – це процес розробки фінансово-

економічних прогнозів, побудований на наукових методах пізнання соціально-економічних явищ з використанням усієї сукупності методів, засобів та способів прогностики.

Ці положення визначають сутність прогнозу, але вони недостатні при оцінці ролі прогнозу в процесі планування. Необхідно акцентувати увагу на тому, що прогноз, по-перше, формує поле об'єктивно досяжних цільових показників розвитку економічної системи, за межі яких не повинні виходити показники плану; по-друге, прогноз цінний тим, що визначає можливі сценарії майбутнього розвитку, порівняльний аналіз яких дозволяє вибрати оптимальний варіант для планування, в якому найкращим чином поєднуються такі критерії: відповідність цілям, ресурсна забезпеченість, ефективність і мінімізація ризиків.

З такого розуміння зв'язку плану і прогнозу випливає ряд вимог до прогнозу.

а) в основі визначення варіантів прогнозу повинні лежати:

- 1) гіпотези про альтернативні моделі поведінки суб'єктів економічних відносин;
- 2) гіпотези про можливі зміни зовнішньополітичних та зовнішньоекономічних умов;
- 3) оцінки можливих соціально-економічних наслідків реалізації кожного варіанту прогнозу;
- 4) оцінки ймовірності ризиків і масштабів наслідків настання ризикових ситуацій для кожного варіанта прогнозу.

б) оцінка рівня достовірності прогнозу (достовірний, щодо достовірний, можливий);

в) прогнозування доцільно організувати як безперервний процес, в ході якого прогнозні параметри коригуються з урахуванням результатів розвитку за минулий період.

Отже, прогнози є одним із основних елементів інформаційного забезпечення планування та управління, зокрема в фінансово-економічній

сфері. Основною метою прогнозування тут є отримання повної інформації, необхідної для передбачення, усвідомлення і своєчасної адаптації цілей і можливостей під обставини. Фінансове прогнозування дозволяє виявити і раціонально використовувати резерви для економічного зростання. Детальніше можна виділити наступні задачі:

- визначення можливих напрямків розвитку суспільства та економічних ресурсів, що забезпечують їх досягнення;
- визначення найбільш імовірних та економічно обґрунтованих варіантів довгострокових та поточних планів;
- отримання науково обґрунтованих варіантів тенденцій розвитку показників якості, елементів витрат та інших показників, які використовуються для розроблення перспективних планів;
- обґрунтування напрямків економічної, технічної та соціальної політики;
- передбачення наслідків рішень та заходів, що приймаються в даний час.

Якісне прогнозування дозволяє виявляти тенденції розвитку ринку і здійснювати свою діяльність відповідно до даних тенденцій, займати лідируючу позицію на ринку і успішно розвиватися, тому держава і великі компанії та середній бізнес витрачають великі кошти на прогнозування процесів.

Актуальність даної тематики пояснюється тим, що прогнозування є основою успішної побудови фінансової та економічної систем. Чим успішніше і точніше побудований прогноз, тим ефективніше буде результат будь-якої сфери діяльності. Однак фінансове прогнозування є, мабуть, найбільш важливим аспектом всієї системи прогнозування.

Таким чином, можемо сказати, що прогноз є пошуком реалістичного, економічно правильного рішення для управління об'єктом чи процесом, а прогнозування є необхідним та важливим науково-аналітичним етапом загального процесу планування.

1.3 Короткий огляд моделей прогнозування

1.3.1 Оптимальний предиктор і його властивості

Для того щоб розробити різні прогнози використовують багато практик, що спираються на цілу систему методів. Всі вони базуються на оцінюванні (прогнозування) величин, що є невідомими для безпосереднього спостереження в даний момент, і їх необхідно знайти за допомогою доступних для вимірювання дотичних величин способів.

При відомій сукупності спостережень побудова функцій від цих величин, які можна буде використати для оцінювання прогнозованого значення є основним завданням стохастичного прогнозування. Назва такої функції - предиктор. Щоб оцінити точність предиктора, часто використовують середньоквадратичну похибку (СКП) хоча можливі і інші підходи.

Оптимальним предиктором, відповідно, зазвичай називають предиктор, який мінімізує СКП в заданому класі. Предиктор також називають прогнозом. Основною ідеєю стохастичного прогнозування є розробка та дослідження методів побудови оптимальних предикторів. На підставі предиктору можна отримати варіанти прогнозу, що відповідатимуть сформульованим гіпотезам та умовам, врахованим під час його побудови.

За супровідними величинами \bar{Y} теорія стохастичного прогнозування величини y_y передбачає, що сумісний закон розподілу $F(\bar{Y}, y_y)$ - відомий. У реальному житті зазвичай все трохи не так і точний вид залежності між y_y та \bar{Y} найчастіше невідомий, тому пошук найкращого предиктора обмежується лише лінійними прогнозами. В таких умовах оптимальний лінійний предиктор існує та задається як умовне математичне сподівання, тобто збігається із функцією регресії y_y на \bar{Y} . Вона безперечно має максимальну кореляцію з y_y серед усіх лінійних предикторів. Очевидно, для

побудови оптимального предиктора досить знати перші та другі моменти початкового розподілу $F(\bar{Y}, y_y)$, які знаходять шляхом оброблення результатів відомих спостережень. Далі використовуючи ці оцінки замість теоретичних характеристик, ми отримуємо емпіричний предиктор, який і використовують для прогнозування майбутніх значень y_y .

Отже, вибір відповідної моделі має головну роль у статистичному підході до прогнозування. Якщо її наповнити числовими параметрами - вона стає безпосереднім інструментом прогнозування - предиктором.

1.3.2 Регресійні моделі

Безперечно вивчення залежності між змінними та визначення кількості факторів є основною задачею регресійного аналізу: рівняння регресії як ніщо точно характеризує кореляційну залежність ендогенної змінної від екзогенних. Сьогодні розрізняють модель (рівняння) парної регресії

$$y_t = a_0 + a_1 x_t + \varepsilon_t \quad (1.1)$$

та множинної регресії

$$y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + \dots + a_m x_{mt} + \varepsilon_t, \quad (1.2)$$

або в матричному вигляді

$$Y = AX + \vec{\varepsilon}, \quad (1.3)$$

де $\vec{\varepsilon}$ – вектор випадкових змінних.

Як бачимо, рівняння (1.2) забезпечує встановлення статистичного взаємозв'язку досліджуваних показників, та в разі його значущості – визначення аналітичних і прогнозованих оцінок. Для оцінки параметрів зазвичай використовують метод найменших квадратів (МНК) за умови мінімуму функціонала

Прогнозування за допомогою регресійних моделей поділяється на наступні етапи:

- в першу чергу це визначення мети дослідження, а також вибір відповідної теорії, що пояснює поведінку системи.
- наступний крок це побудова системи показників та параметрів, які мають значний вплив на кожен конкретний показник, та розробка схеми прогнозу.
- вибір способу зв'язку показників між собою та відібраними чинниками;
- побудова економетричної моделі, тобто відображення теорії у вигляді рівняння або системи рівнянь регресії і тотожностей, яка пов'язує зібрані змінні. Особливу увагу тут варто надати змінним, що містять інформацію про перспективу на майбутнє, а також випередженню та запізнення впливу змінних у рівняннях;
- пошук даних про значення змінних (варто дотримуватись теоретичних правил);
- аналіз інформації;
- оцінювання невідомих параметрів рівняння моделі використаного методу;
- впевнюємось, що побудована модель адекватна, точна і має хорошу прогнозовану спроможність; а також на безпосередньо відповідає досліджуваному економічному процесу;
- використання знайденої найкращої моделі для прогнозування.

За допомогою отриманих рівнянь з оціненими параметрами та прогнозованими змінними, зазвичай роблять передбачення потрібних показників, а саме значень ендогенних змінних. Для того щоб спрогнозувати кілька кроків чи періодів вперед, використовують просто послідовні прогнози на один період. Використовуючи інші джерела, наприклад, іншу економетричну модель або експертні методи або одновимірні моделі часових рядів можна знайти значення величин екзогенних змінних, від яких суттєво залежить прогноз.

Визначимо переваги регресійних моделей:

- аналітично відображують взаємозв'язки між показником та факторами, що досліджуються;
- дають можливість оцінити міру впливу окремих факторів на показник що є дуже цінною характеристикою;
- забезпечують визначення оцінки впливу всіх факторів на показник;
- дають можливість отримати вірогідні результати прогнозування навіть при складному зв'язку між змінними навіть для складної динаміки розвитку об'єкта;
- вони підлягають перевірці на адекватність моделі сучасними математичними методами; це під сумнівом
- одне з найбільш приємних - є достатньо простими в реалізації.

1.3.3 Прогнозування за допомогою нейронних мереж

Усім відомо наскільки потужним інструментом прогнозування сьогодні є нейронні мережі (НМ). Вони принципово відрізняються від звичного нам способу програмування та аналізу результатів. Нейронні мережі потрібно навчати, і це одна з найбільших їх переваг в порівнянні з іншими відомими нам алгоритмами.

При цьому для визначення того, що потрібно прогнозувати, необхідно вказувати змінні, які використовуються і передбачаються. На цьому етапі варто звернути більше уваги на потрібний рівень деталізації. Він залежить від багатьох факторів: доступність необхідних даних та їх точність, вартість проведення аналізу і навіть аналіз переваг при отриманні результатів прогнозування мають значення. Часто виникають ситуації коли не можна однозначно вибрати найкращий набір змінних, тоді варто пробувати різні альтернативи та вибрати один з варіантів з найкращими результатами або ж зкомбінувати отримані результати (про це далі детальніше). Зазвичай це основний спосіб вибору кращого результату при розробці прогнозуючих систем, що засновані на аналізі історичних даних.

Також при побудові нейромереж велику роль відіграє визначення: періоду прогнозування (одиниця часу, на яку робиться прогноз), горизонту прогнозування (число періодів) та інтервалу прогнозування (частота прогнозу). Також часто інтервал прогнозування може співпадати з періодом прогнозування, це залежить від предметної області дослідження. Для фінансово-економічної сфери період і горизонт прогнозування зазвичай, доба. Наразі відомі три способи навчання нейронних мереж:

- ‘з вчителем’ - нейромережа знає правильні відповіді (результат) на кожен приклад навчальної вибірки. Ваги налаштовуються так, щоб

мережа подавала відповіді найбільш близькі до правильних відповідей - саме такий підхід використовується в роботі;

– ‘без вчителя’ (самонавчання) - не вимагає знання правильних відповідей на кожен приклад. У цьому випадку важлива внутрішня структура даних та залежність між даними в навчальній множині, що дозволяє розподілити приклади по категоріям;

– змішана - відповідно частина ваг розраховується за допомогою навчання із вчителем, інша - без.

Алгоритм навчання нейронної мережі включає в себе:

- отримання тимчасового ряду з інтервалом;
- заповнення «прогалин» в даних;
- згладжування ряду будь- яким методом (найпопулярніший - метод ковзних середніх);
- отримання ряду відносної зміни прогнозованої величини;
- створення таблиці «вікон» з глибиною занурення тимчасових інтервалів;
- заповнення таблиці додаковими данми (наприклад, зміна величини за попередні роки тощо);
- шкалювання результатів;
- визначення навчальної та валідаційної вибірок;
- підбір необхідних параметрів нейромережі;
- її навчання;
- Перевірка працездатності та якості нейромережі на перевірочній вибірці.

1.3.4 Лінійні параметричні моделі

Стаціонарні часові ряди можна представити широким класом лінійних параметричних моделей. Найбільш поширеними є моделі авторегресії (АР), ковзної середньої (КС) та змішані (АРКС). Спектр застосування цих моделей не обмежується стаціонарними процесами - так, ряди зі специфічною однорідною нестаціонарністю можна звести до стаціонарних і описувати модифікованою формою моделі АРКС, відомої як модель Бокса-Дженкінса. Лінійні параметричні моделі мають загальну назву авторегресійні інтегровані моделі ковзної середньої (АРІКС). Вони ґрунтуються на припущенні лінійності процесу породження даних і описують стаціонарний процес, який має три ознаки: p — порядок авторегресії, d — необхідний порядок інтегрування, тобто кількість разів взяття різниць для зведення початкового часового ряду до стаціонарного, q — порядок ковзної середньої в моделі. Поєднання різних моделей аналізу часових рядів у межах однієї дає змогу працювати з моделями невисоких порядків, що суттєво розширює сферу практичного застосування їх. Окрім того, з'являється можливість розробляти модель за допомогою однакових статистичних характеристик — автокореляційних і часткових автокореляційних функцій, розробляти спільний алгоритм для обчислення параметрів моделі, однаковим чином будувати прогноз на підставі побудованої моделі тощо.

Будь-які різновиди АРІКС-моделей є окремим випадком загальної лінійної моделі часового ряду, яка є базовою для теоретичних досліджень стаціонарних процесів. В основі її визначення лежить поняття «білого шуму».

Якщо вважати, що середнє значення стаціонарного ряду дорівнює нулю, то загальна лінійна модель це стаціонарний процес у вигляді лінійної комбінації білого шуму з різними ваговими коефіцієнтами.

В АРІКС-моделях, під час прогнозування змінної для майбутнього моменту часу, лагові значення цієї змінної, які слугують пояснюючими змінними (регресорами) моделі, можна розглядати або фіксованими на вибіркових значеннях, або випадковими. Перша можливість призводить до умовного прогнозу, на кшталт множинної регресії, друга — до безумовного прогнозу. Отже, у прогнозуванні за моделлю типу АРІКС розглядають умовні та безумовні прогнози. Відомо, що умовна дисперсія випадкової величини не перевищує її безумовну дисперсію, тому точність умовного прогнозу завжди вища.

Якщо модель правильно специфіковано, то можливі два джерела помилок прогнозів: невизначеність майбутніх значень випадкової величини ε , відсутність точних значень коефіцієнтів моделі (наявні тільки їхні оцінки). Під час прогнозування за моделлю АРІКС від наявної вибірки залежать як оцінки коефіцієнтів моделі, так і значення регресорів, тому важко аналітично виразити умовну дисперсію помилки прогнозу через спостереження часового ряду. Як правило, обмежуються припущенням про те, що коефіцієнти відомі точно. Зрозуміло, що таке припущення зменшує дисперсію помилки прогнозу й тим самим збільшує умовну точність як умовного, так і безумовного прогнозів. Для досягнення мінімуму середньоквадратичної помилки (MSE) потрібно обчислити умовне математичне сподівання.

1.4 Критерії якості моделей та оцінок прогнозів

Сьогодні можна знайти велику кількість критеріїв, які використовуються для визначення якості побудованих моделей прогнозування, їх точності та адекватності.

Що називають ефективністю моделі? Фактично це набір кількісних показників, які дозволяють описати найбільш адекватну та оптимальну за своєю структурою модель.

Найвідомішими критеріями ефективності сьогодні є: середньоквадратична похибка (mean squared error, MSE) та сума квадратів похибки (sum squared error, SSE), вони розраховуються за наступними формулами:

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2, \quad (1.4)$$

$$SSE = \sum_{i=1}^n e_i^2. \quad (1.5)$$

Найчастіше ці критерії використовуються в пакетах прикладних програм для прогнозування при виборі оптимальної моделі прогнозування [15]. Для того щоб підкреслити найбільші похибки прогнозу, кожне значення помилки підноситься до квадрату, на відміну від таких критеріїв, як середнє абсолютне відхилення. Варто звернути увагу на таку особливість, так як модель прогнозування, яка постійно дає середні по величині похибки, в деяких випадках може бути більш раціональною за іншу модель, що має малі відхилення, але періодично дає неприйнятні за величиною викиди.

У випадках, коли потрібно провести порівняння якості прогнозування різномірних об'єктів доцільно використовувати середню абсолютну похибку у відсотках (mean absolute percentage error, MAPE) так як вона не залежить від масштабу даних. Це є значною перевагою моделі. Оцінка можна порахувати за формулою

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|e_i|}{Y_i} * 100 \quad (1.6)$$

де n – розмір ряду; e_i – похибка прогнозу; Y_i – фактичне значення ряду в момент i . З цією величини бачимо наскільки у порівнянні із дійсним значенням ряду є великою похибка.

Звернемо увагу коефіцієнт детермінації, який розраховується за формулою

$$R^2 = 1 - \frac{SSE}{\sum_{i=1}^n (Y_i - \underline{Y})^2} \quad (1.7)$$

де n – розмір ряду; Y_i – фактичне значення ряду в момент i ; $\underline{Y} = \frac{1}{n} \sum Y_i$ – середнє значення ряду. Для адекватної моделі повинно виконуватись $R^2 \rightarrow 1$. На практиці прийнято вважати адекватною модель якщо $R^2 > 0.75$ [3].

Звичайно як і у всіх моделей, в вищеописаних також є недоліки. Тут при ускладненні моделі їхні значення часто свідчать про покращення точності моделі, хоча це часто зовсім не означає покращення якості моделі. Проте цю особливість можна взяти до уваги за допомогою скорегованого коефіцієнту детермінації [12]

$$R_{adj}^2 = 1 - \frac{n-1}{n-r-1} (1 - R^2) \quad (1.8)$$

де r – кількість параметрів конкретної моделі.

Проте, навіть близькі до одиниці значення R^2 та R_{adj}^2 можуть не свідчити про хороші прогнозні якості моделі: тут необхідною є відсутність автокореляції похибок, яку можна перевірити за допомогою статистики Дарбіна-Уотсона. Але якщо в моделі серед пояснюючих змінних наявні лагові значення результативної ознаки статистику застосувати неможливо. В таких випадках використовують Q-статистику Льюнга-Бокса (Ljung-Box Q-statistic, LBQ), яка розподілена за законом χ^2 і розраховується за формулою

$$Q_{LB} = N(N + 2) \sum_{i=1}^s \frac{r_k^2}{N-k} \quad (1.7)$$

де N – розмір ряду; r_k – автокореляція k -го порядку; s – кількість лагів, що перевіряються. Ідеєю моделі є перевірка відмінності від нуля відразу декількох коефіцієнтів автокореляції.

Якщо $Q_{LB} > \chi_{1-\alpha,s}^2$, де $\chi_{1-\alpha,s}^2$ – квантилі розподілу χ^2 з m ступенями вільності, то тоді визнається наявність в часовому ряді автокореляції до s -го порядку. Ця статистика придатна як для великих так і для малих кількостей даних тому що побудована на основі статистики Бокса-Пірса. Слід зазначити, що статистику Льюнга-Бокса застосовують до залишків, а не початкового ряду на відміну від статистики Бокса-Пірса.

Далі варто приділити увагу критеріям якості моделей, які базуються на системі ‘штрафів’ через складність такої моделі. Основною відмінністю між такими критеріями є міра «жорсткості», тобто величина штрафу за більшу кількість параметрів в моделі.

Розглянемо критерій Малоуза:

$$C_p = \frac{SSE}{\sum_{i=1}^n (Y_i - \underline{Y})^2} + \frac{2(r+1)}{n-r-1} \quad (1.8)$$

та інформаційний критерій Акайке:

$$AIC = \ln(MSE) + \frac{2(r+1)}{n} \quad (1.9)$$

Байєсівський інформаційний критерій (критерій Шварца):

$$BIC = \ln \ln (MSE) + \frac{(r+1) \ln \ln n}{n} \quad (1.10)$$

При мінімізації цих критеріїв якість моделі значно зростає. Також позитивно впливає на описані вище критерії зменшення залишкової дисперсії та кількість врахованих параметрів.

При великих n критерії C_p та AIC сильно корелюють між собою. Проте, як можна помітити з наведених формул BIC є найбільш жорстким критерієм, причому, його жорсткість зростає з ростом n , на відміну від інших критеріїв. Критерії C_p та AIC безпосередньо направлені на досягнення високої якості прогнозу: C_p направлений на мінімізацію дисперсії похибки прогнозу, а AIC – на мінімізацію розходження між щільністю розподілу по істинній моделі та вибраній. Вимога максимізації ймовірності вибору істинної моделі покладено в основі BIC .

Зазначимо також, що критерій Малроуза можна використовувати лише при умові нормального розподілу залишків. І також варто відмітити значний недолік критерію Акайке, який полягає в переоцінці істинної кількості параметрів в моделі.

Для визначення адекватності моделі прогнозування використовують коефіцієнт неспівпадіння Тейла:

$$U = \frac{\sqrt{\sum_{t=1}^n (\hat{y}_t - y_t)^2}}{\sqrt{\sum_{t=1}^n \hat{y}_t^2 + \sum_{t=1}^n y_t^2}} \quad (1.11)$$

де \hat{y}_t – прогнозні значення часового ряду; y_t – реальні значення часового ряду; t – момент часу. Прогноз прийняти вважати точним за умови $U \rightarrow 0$.

1.4 Підхід до прогнозування з використанням комбінованих оцінок

На сьогодні найбільш відомими та поширеними способами комбінування оцінок прогнозів є:

- усереднення прогнозів (вагові коефіцієнти при цьому однакові);
- зважене усереднення прогнозів.

Розглянемо детальніше перший варіант. Комбінований прогноз визначається цим способом досить просто:

$$\hat{y}(k) = \sum_{i=1}^n \hat{y}_i(k) \quad (1.12)$$

де $\hat{y}(k)$ - комбінований прогноз; $\hat{y}_i(k)$ - прогнози, отримані різними методами. У випадку незміщеності прогнозів (це безпосередньо має забезпечувати метод прогнозування), комбінований прогноз також обов'язково буде незміщеним. Похибка комбінованого прогнозу (лише з метою спрощення розглянемо варіант комбінування двох значень для обчислень):

$$e(k) = y(k) - \hat{y}(k) = y(k) - \hat{y}_1(k) + \hat{y}_2(k) \quad (1.13)$$

де $y(k)$ – фактичне значення прогнозованої змінної.

Далі пропоную розглянути інший спосіб усереднення оцінок прогнозів - зважене середнє. Якщо у нас немає інформації щодо характеристик індивідуальних прогнозів, то можна присвоїти різні вагові коефіцієнти окремим прогнозам на основі суб'єктивних або ж експертних суджень:

$$\hat{y}(k) = \sum w_i \hat{y}_i(k) \quad (1.14)$$

де w_i – вагові коефіцієнти; $\sum w_i = 1$. Очевидно, що більші значення вагових коефіцієнтів необхідно присвоювати тим індивідуальним прогнозам, які мають меншу дисперсію похибок.

Як же правильно вибрати найкращий коефіцієнт? зазвичай, похибки прогнозів для конкретних моделей і процесів відомі, або їх можна визначити. Це дає нам можливість конструктивно підійти до розв’язку задачі вибору вагових всіх коефіцієнтів. Моделі, які дають менші суми квадратів похибок прогнозів, очевидно, генерують якісніші прогнози, то цілком зрозуміло приймемо цю міру за основу для визначення вагових коефіцієнтів. Позначимо суму квадратів похибок прогнозування (для історичного прогнозу) через

$$sse = \sum e^2(k) \quad (1.15)$$

де sse – суми квадратів похибок для кожного з методів, що використовуються в нашому випадку [1].

1.5 Висновки до розділу та постановка задачі

На мою думку, сьогодні питання якісного регулювального підходу до прогнозування та управління фінансово-економічними процесами залишається відкритим. Що допоможе урегулювати фінансово-економічні процеси на рівні держави, як запобігти різноманітним економічним кризам, що виникають сьогодні навіть у розвинутих країнах? Коли економічні процеси та фінанси починають різко змінювати свій тренд, або їх рух стає неконтрольованим? Економічні теорії формують деякі ідеальні моделі, від яких реальні фінансово-економічні процеси, звичайно, достатньо далекі. Коректне розв’язання задач прогнозування – актуальне завдання державного управління, оскільки прогнози розвитку процесів складають основу для їх

дальшого планування та управління. Це особливо стосується фінансово-економічних процесів, що мають місце в макроекономіці. На мою думку, побудова якісних математичних моделей дозволить глибше зрозуміти сутність необхідних явищ, обґрунтувати вибір необхідних інструментів регулювання для динаміки фінансових та економічних процесів навіть на державному рівні [2, 3].

Сьогодні головною задачею що може бути розв'язана за допомогою прийняття економічних чи фінансових рішень орієнтованих у майбутнє є зниження неминучої невизначеності. Для цього фінансовий та економічний аналіз може використовуватися як один із інструментів обґрунтування короткострокових чи довгострокових рішень, доцільності інвестицій; а також як спосіб оцінки якості управління та правильності прийнятих рішень в минулому; або ж як спосіб прогнозування очікуваних фінансових чи економічних результатів. За допомогою такого прогнозування можливо значно мірою поліпшити управління компанією (можливо навіть державою!) за рахунок забезпечення координування всіх факторів виробництва і реалізації між собою, взаємозв'язку діяльності всіх підрозділів, і безпосередньо розподілу спільної відповідальності.

Станом на сьогодні існує багато статистичних та ймовірнісних методів для розв'язку такої задачі прогнозування, серед яких найчастіше застосовуються моделі на основі авторегресії, експоненційного згладжування, нейромереж, а також все більшої популярності набувають мережі Байєса.

Суттєвим недоліком авторегресійних моделей є велика кількість вільних параметрів, що вимагають ідентифікації. Недоліком же непромережених моделей є їх непрозорість моделювання та складність навчання мережі. Просте кожен з методів мають свої значні переваги для застосування в тому чи іншому випадку, чим не можна нехтувати.

Роботи дослідників останніх років свідчать про різке поширення застосування мереж Байєса сьогодні. Побудова мережі тісно пов'язана з необхідністю розв'язання кількох задач обчислювального характеру, що

зустрічаються при навчанні мережі. Це ускладнює використання методів для задач з багатьма параметрами та великою кількістю даних/

Сьогодні безумовно актуальними залишаються питання покращення якості прогнозів існуючих моделей та створення адаптивних (комбінованих) моделей на основі вже існуючих.

Основною задачею роботи є виконання огляду сучасних методів короткострокового прогнозування, збір необхідних статистичних даних для побудови моделей і оцінювання прогнозів; розробка програмного продукту для виконання обчислювальних експериментів з можливістю його подальших модифікацій для запуску стартап-проекту. Побудова математичної моделі для обраних процесів, пошук функцій прогнозування, обчислення оцінок прогнозів; виконання комбінування цих оцінок для отримання найточніших результатів прогнозування для різних процесів та вибору кращого способу комбінування.

РОЗДІЛ 2

ВИБІР ТА ОПИС МЕТОДІВ СТАТИСТИЧНО-ЙМОВІРІСНОГО ПРОГНОЗУВАННЯ

2.1. Мережа Байєса

Останнім часом мережі Байєса почали набирати популярність серед способів моделювання процесів різної природи та стали окремим напрямом в області інтелектуального аналізу даних. Мережі Байєса широко застосовуються для обробки часових рядів (статистичні дані), та інших все можливих типів представлення даних, наприклад, експертні оцінки, лінгвістичні змінні тощо. [11].

Окремо варто відмітити, що байєсовські мережі знайшли застосування у розв'язанні задач медичної діагностики. Відмічається великий прорив в уточненні діагнозів найрізноманітніших хвороб при умовах неточної та неповної інформації. Тільки через цю їх особливість, вже варто приділяти значну увагу їх вивченню та розвитку, на мою думку. Також мережі популярні в системах класифікації даних різної природи, системах автоматичного розпізнавання мовних сигналів, маркетингу і бізнесі, та у багатьох інших сферах людської діяльності. Також мережі байєса допомагають забезпечити можливість прогнозування модифікацій запитів в пошукових системах, які надходять від користувачів та для діагностики проблем у різних системах типу “wizard” та навіть принтерах. Одним з найпопулярніших застосувань наївного баєсового алгоритму, з яким, я впевнений, кожен з нас зустрічався не один раз є фільтрація поштового спаму. В підсумку мережа допомагає визначити причинно-наслідкові зв'язки між подіями, допомагає встановити ймовірність настання тієї чи іншої події при отриманні нової інформації щодо зміни стану будь-якого вузла (змінної) мережі. Проте наскільки успішним буде використання цього методу моделювання а також створення статистичного висновку залежить в багато чому від досвідченості експерта у вмінні чітко

сформулювати умову задачі, обрати залежні змінні процесу щоб в достатній мірі описати його статику чи динаміку, правильно використати і сформувати вибірки для навчання мережі (поділивши на навчальну та перевірочну), а також вміти коректно сформувати результат – висновок за допомогою побудованої мережі [13].

Можливість врахування водночас кількісних та якісних показників, правильно використати надходження нової інформації динамічно, використання явної залежності між існуючими факторами (що в свою чергу впливають на фінансові показники) є, мабуть, основною перевагою мережі байєса. Сюди ж можна віднести можливість врахування у моделі дискретних і неперервних змінних, врахування невизначеностей та практично відсутнє обмеження на кількість змінних, а також, звичайно, наочність моделювання . На відміну від інших моделей на основі регресійних рівнянь, мережні не вимагають повноти інформації тому опис мережі зазвичай не є складним. Обмежень на закони розподілу змінних також немає, що дає водночас гарні показники як інструмент для класифікації даних. І як бонус, можна додати що мережі байєса є дуже зручними для аналізу та користувачів в цілому так як мають дуже вдале представлення у вигляді графів.

Проаналізувавши найпопулярніші методи інтелектуального аналізу даних можна судити, що мережі байєса у порівнянні з відомими моделями “black box”, демонструють логічну модифікацію та інтерпретацію структури взаємозв’язків між змінними у явній формі, дають змогу враховувати попередній апріорний досвід експертів і, безумовно, надають зрозуміліше пояснення своїх висновків. Мережі є успішними при розв’язанні реальних задач так як ґрунтуються на фундаментальних законах, аксіомах, тверджень і положень теорії ймовірностей, які досліджувались та вдосконалювались сотнями років.

Перейдемо до недоліків методу. При дослідження можливих застосувань методу та його досліджень було знайдено ряд недоліків, пропоную розглянути їх детальніше далі. При використанні мережі байєса

часто виникають труднощі при знаходженні методу точного імовірнісного висновку базуючись на навчальній вибірці: для того щоб порахувати ймовірності вершин вузлів, потрібно використати матрицю спільних ймовірностей замість значень умовних ймовірностей емпіричних значень. Основною метою є те, що це дозволяє не розробляти структуру мережі заздалегідь. Також мережу просто реалізувати на при цьому досягти бажаної обчислювальної швидкості (залежить від розміру вибірки)[13; 14]. також не можна нехтувати складністю(чи навіть проблемою) вибору змінних для побудови моделі. Для цього використовується алгоритм конструктивного вибору змінних для створення високопродуктивної мережі байєса для прогнозування.

Мережі Байєса відносяться до імовірнісних мереж, які за допомогою формули Байєса та її узагальнень дізнаються необхідні знання про ймовірності у вершинах мережі. Це по своїй суті графічні моделі подій і процесів на основі об'єднання результатів дослідження теорії ймовірностей та аналізу теорії графів. Для того щоб побудувати графічну модель застосовують поняття “модульність” що являє собою поділ великої складної системи на прості елементи. Далі ж, для об'єднання окремих елементів назад в одне ціле використовують отримані результати теорії ймовірностей, які забезпечують моделі хорошу дієздатність для практичних задач, а також навіть дозволяють об'єднувати бази даних та графічні зображення. Такий підхід до побудови моделі- з графами та теорією дає аналітику поштовх для створення структури даних для наступної розробки, винайдення ефективних алгоритмів обробки даних та прийняття рішень як результат, а також дає можливість будувати моделі процесів навіть при наборі змінних які сильно взаємодіють між собою [11].

Якщо розглядати мережу байєса із математичної точки зору, то можна сказати що це модель подання існуючих і відсутніх імовірнісних залежностей. При цьому зв'язок $A \rightarrow B$ називають причинним, якщо подія A – причина виникнення B , тобто коли існує механізм, при якому значення, прийняте A ,

безпосередньо чи ні впливає на значення, прийняте В. Коли всі зв'язки мережі являються причинними, тоді таку мережу називають каузальною.

Для полегшення формування байєсівського висновку на відомій множині конкуруючих гіпотез, які пояснюють множину даних, варто описати необхідну послідовність дій (алгоритм). Отже, для кожної гіпотези необхідно виконати наступне:

- правдоподібну чи апріорну інформацію, що міститься в даних, у ймовірності записати як ймовірності;
- обчислити добуток ймовірностей;
- з метою отримання апостеріорної ймовірності для кожної гіпотези-нормувати результати (при наявній інформації)
- обрати гіпотезу, яка має максимальну ймовірність.

Деколи статистичних даних достатньо щоб обчислити апріорні ймовірності, проте, в більшості випадків це неможливо зробити через суб'єктивну складність отримання статистичних даних, а апріорні знання можуть представлені у інших формах [13; 15].

Наведемо приклад мережі Байєса на рисунку 2.1 нижче.

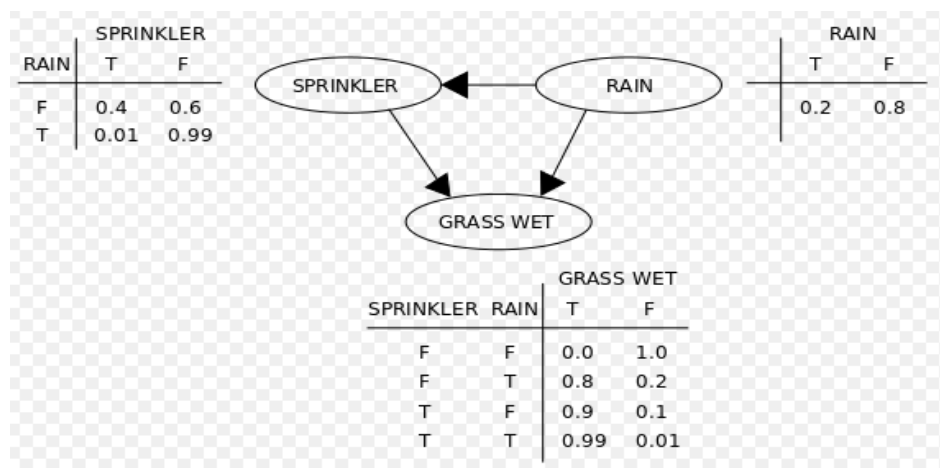


Рисунок 2.1 - Приклад мережі Байєса

2.2 Прогнозування за допомогою нейронних мереж

Як вже було сказано й доведено вище в роботі, задачі прогнозування на основі моделей часових рядів сьогодні присвячено багато досліджень. Значних обертів та поширення набувають побудови прогнозу на основі використання ймовірнісних методів та суб'єктивних знань аналітиків. Зупинимось ще раз на недоліках, які мають ймовірнісні методи:

- відсутність у моделі поняття структури та системи зв'язків реального об'єкта;
- складність побудови моделей при умові, що дані мають тимчасові зрушення відносно один одного та зберігаються в різних часових рядах;
- часто можлива недостатня точність прогнозу;
- велика чутливість отриманих значень до недостатньої інформації та (або) її зашумленість;
- висока залежність якості отриманого прогнозу від кваліфікації та досвідченості експерта в конкретній предметній області.

Таким чином, зазначені недоліки підходу приводять до необхідності розробки нових математичних моделей на основі методів штучного інтелекту а також алгоритмів і спеціалізованого програмного для забезпечення вирішення поставлених задач, що підвищують надійність і точність прогнозів. Також вони здатні працювати в умовах недостатньої інформації або навіть її значної зашумленості, що дозволяє отримати результат за короткий час. Це є особливо цінним для задач оперативного керування.

Нейронні мережі є одним із найпопулярніших методів вирішення задачі прогнозування часових рядів математичних моделей, вони включають в себе розвинену методологію структурного моделювання й методів навчання, що базуються на добре розвинутій теорії нелінійного програмування. Метою прогнозування є зменшення ризику при прийнятті рішень. Прогноз звичайно

виходить помилковим, але помилка залежить від прогнозуючої системи, яка використовується.

Безумовно програмні реалізації розглянутих алгоритмів повинні мати дружній інтерфейс для користувача, дозволяти фахівцеві конкретної предметної області моделювати процес самостійно, без участі фахівця-математика.

За допомогою збільшення ресурсів для прогнозування ми маємо можливість покращити точність та навіть зменшити втрати що пов'язані з невизначеностями при процесі прийняття рішень. Не потрібно забувати, що ніякий метод прогнозування ніколи не зможе повністю нівелювати ризик при прийнятті рішень, тому ми завжди повинні явно визначати неточність прогнозу. Тому остаточне рішення повинне базуватись на результаті отриманого точного прогнозу з врахованою можливою помилкою прогнозування. Іншим перспективним кількісним методом прогнозування окрім мереж байєса які описані вище, є використання нейронні мережі, дослідження яких набирає шалених обертів останні роки. У них є велика кількість переваг, основні з яких пропоную розглянути далі.

Першою перевагою безумовно є гнучка структура. В межах обраної архітектури мережі експерту варто лише налаштовувати та регулювати кількість шарів для навчання та нейронів, чим можна змінити (оптимізувати як ціль) результати глобально. Також можна підібрати іншу активаційну функцію, що додає простору для маневрів. Такі налаштування дозволять підібрати найкращий стан мережі при обраній архітектурі та максимізувати точність навчання мережі, а отже отримати найкращий прогноз та мінімальну похибку. Маючи інформацію про минулі події для вибірок навчання такі кроки та оптимізація мережі працюватимуть навіть для найпростішої архітектури - одношарового персептрону. При цьому можна навіть змінювати видимість зовнішніх параметрів прост включаючи\ виключаючи відповідний нейрон мережі.

Наступна перевага - адаптивність, тобто незалежність експерта від математичної моделі поведінки досліджуваного часового ряду, адже власне процес побудови нейромережевої моделі проходить під час навчання без додаткових зусиль зі сторони експерта. Як саме це відбувається? експерт надає мережі надаються дані навчальної вибірки з бази, а далі вона сама налагоджується під них.

Основний недолік - недетермінованість нейромереж. Тобто неможливість дізнатись логіку прийняття рішень мережею, така поведінка називається «чорним ящиком», який якийсь працює, але логіка прийняття рішень прихована від людини. Щоб створити на основі мережі експертну систему були придумані алгоритми витягу знань із нейронної мережі, які допомагають розбити та формалізувати навчену нейронну мережу на список логічних правил [20]. Відомими вченими які займались дослідженнями у області штучних нейронних мереж є зокрема такі вчені, як Бодянський Є.В., Руденко О.Г., Галушкін А.І., Барт Коско, Козадаєв А.С. та багато інших [12]. Судячи з популярності методу, сподіваюсь з'явиться багато молодих експертів та нових відкриттів.

Розглянемо основні етапи створення та прогнозування за допомогою нейронних мереж. Підготовчий етап - це збір даних, що включає пошук високоякісних даних з доступних джерел, які містять повну й достатню інформацію в необхідній області та їх поверхневий аналіз на предмет доцільності використання. В цілому на цьому етапі потрібно врахувати наступне:

- доцільність знайдених даних та її доступність в часі (велику за обсягом чи періодичну інформацію часто буває важко отримати. Історичні дані значно замінюються, наприклад дані за день потім складаються в тиждень, місяць т.д);

- способи відновлення пропусків в спостереженнях (Найбільш відомими для відновлення пропущених спостережень є два підходи: перший полягає в тому, щоб просто виключити рядок або стовпець матриці даних, де є

пропущені значення. Проте тоді часто такі прогалини не дозволяють побудувати рівний часовий ряд. За іншою альтернативою можна припустити, що пропущені дані в середньому мають характер, аналогічний наявним, і можуть бути відновлені як, наприклад, середні значення із кількох сусідніх елементів даних);

– очищення та фільтрація даних (очевидно, що зашумлені дані значно спотворюють процедуру оцінки та негативно впливає на результат прогнозу. Очищення даних повинне передбачати нівеляцію або видалення небажаних шумових ефектів для одержання більш якісних результатів спостережень для навчання мережі. Крім того, при очищенні даних необхідно порівнювати вхідні значення з передбачуваним діапазоном змін результатів спостережень, для того щоб відділити вихідні дані від тих які входять в межі діапазону викидів);

Як бачимо, перший крок є одним з найвідповідальніших. Навіть при професійно створеній архітектурі мережі, неякісні дані можуть нівелювати всі зусилля і бажаний результат так і не буде досягнутий. Аналіз вхідних даних напряду залежить від задачі яку потрібно вирішити. Саме зібрані для аналізу спостереження представляють собою набір багатовимірних векторів, що зображують досліджуваний процес.

Далі поговоримо про перетворення даних і для чого це потрібно. Найчастіше перетворення необхідне для :

- класифікації вхідних даних;
- трансформації від лінійних задач до нелінійних;
- дослідження в межах лише певного діапазону спостережень.

Найпоширенішими та найпростішими методами є логарифмування та взяття відсоткової різниці які можна легко порахувати в табличному вигляді, складнішими та менш застосовуваними є статистичні перетворення, такі як, наприклад, нахил лінії регресії, що вимірює напрямки трендів у потоці даних, або ще більше складні базуються на перетвореннях Фур'є, які іноді навіть

несуть більше інформації, чим вихідні дані. Проте тут не обійтись без достатньо обширних знань математичного аналізу.

Варто відмітити погану роботу нейронних мереж при вхідних величинах великого різноманітного набору значень в спостереженнях. Для того щоб мінімізувати такий негативний вплив прийнято масштабувати вхідні дані у діапазоні $[0...1]$ або $[-1...1]$. Проте зазвичай функції і методи, що використовуються для масштабування не дають можливості одержання близьких до країв відрізків (0 або 1 або -1 у деяких випадках) вихідних значень. Тому як результат часто можна зустріти масштабування лише в діапазоні $[0,2...0,8]$ або $[-0,8...0,8]$.

Наступна процедура - відбір змінних, вона має на меті знизити розмірність вектору вхідних даних. Це пов'язано з тим, що зазвичай менша кількість вхідних змінних обіцяє більшу ефективність у роботі нейронної мережі, аніж більше число входів нейронні мережі. На сьогодні використовують три основні способи пошуку оптимальної кількості вхідних вузлів мережі. В першу чергу звернемо увагу на використання статистичних методів або методів нелінійної динаміки для вирішення проблем з частотою відліків - перший спосіб. Відомий метод аналізу чутливості для визначення важливості входів може бути застосований як інший спосіб для навчання мережі. Остання процедура базується на ранжуванні вузлів і виключення точок які отримали найнижчі ранги вважається досить ефективною. Це відбувається наступним чином: послідовний відбір множини підготовлених спостережень для входу, далі відбувається навчання мережі при визначеному прийнятому рівні помилки, потім проведення аналізу чутливості та останнім кроком виключення всіх змінних, що перебувають нижче обраного рівня.

Визначення відносного впливу обраного вузла на середнє значення по мережі до власне навчання мережі є відомою альтернативою аналізу чутливості: власне, мережа не може навчитися сотням різних пропонованих їй образів маючи обмежену кількість нейронів у схованих шарах. Як можна оцінити якість та точність розробленої нейронної мережі?

В першу чергу необхідно звернути увагу на наступні показники, вони є найвпливовішими:

- ефективність обраної моделі: щоб порахувати ефективність моделі потрібно спочатку порахувати оцінку поведінки мережі при необхідному або ідеальному виході та порівняти результат з тим, який отримали на виході мережі при подачі на вхід всі елементи навчальної вибірки;

- ефективність роботи нейромережі: ефективність нейромережі є мірою потенціалу роботи навченої мережі, вона визначає якість навчання мережі, базуючись на певній характеристиці, що досягається при необхідному або ідеальному виході мережі. Цей показник оцінює дії нейронної мережі на тестовій вибірці та після цього порівнює отриману оцінку з тією, котра досягається при використанні необхідного (ідеального) виходу;

- її стійкість та стабільність: стійкість характеризує навченої мережі в різні часові періоди, це наприклад можуть бути різні проміжки часу, що перетинаються або не перетинаються, чи змінюють розміри. Такий підхід є популярним так як використання так званих “ковзних вікон” дає розподіл значень характеристики нейронної мережі. Стабільність є найвідомішою характеристикою нейронної мережі, що базується на вікнах змінних розмірів. Для цього при оцінці характеристики навченої нейронної мережі на багатьох тимчасових проміжках, кожний з яких має випадкові дати початку й закінчення все зводиться до використання методу Монте-Карло.

На сьогодні вже створено безліч розробок для всеможливих нейронних мереж, з вбудованими алгоритмами підготовки даних для навчання, поділу вибірок, оптимізації та обрахунку похибки обчислень. Проте навіть не зважаючи на це найбільш популярними всеодно є програмні реалізації різних нейромережових парадигм, які ще називають нейропакетами або нейроеммуляторами [10]. На сьогодні вже існує кілька десятків фірм-виробників нейропакетів, а запропонований ними асортимент розроблених продуктів становить кілька сотень найменувань. Універсальність того або іншого нейропакета є дуже суб'єктивним фактором, тому що той самий

нейропакет може бути й універсальним і спеціалізованим, залежно від того, для рішення яких саме прикладних задач він планується використовуватись. Звісно що повністю універсальних нейропакетів не існує, оскільки неможливо заздалегідь передбачити всі нейронні структури, які можуть знадобитися для рішення тих або інших прикладних задач. Тому при виборі кращого в тому чи іншому випадку нейропакета варто звернути увагу на спектр задач які ним обіцяють бути вирішені найкраще.

Найбільш відомі та розповсюджені :

- NeuroSolutions – фірми NeuroDimension, Inc.;
- Statistica з модулем Neural Networks – фірми StatSoft.;
- Deductor – фірми BaseGroup.;
- NeuroShell2 – фірми Ward Systems Group;
- BrainMaker Pro – фірми California Scientific Software.

Не будемо забувати, що окрім універсальності нейропакет також повинен бути простим у використанні, мати інтуїтивно зрозумілий інтерфейс і забезпечувати наочність подання інформації.

2.3 Метод експоненційного згладжування

Вже в другій половині минулого століття почався активний розвиток методів експоненційного згладжування. Тоді потреба в них виникла для вирішення задачі прогнозування одночасного руху великої кількості різноманітних товарів на складах. І саме тоді ці методи набрали популярності (та власне стали найбільш поширеними) завдяки їх беззаперечній простоті та ефективності. Самими відомими є методи звичайного, подвійного та потрійного згладжування, а також метод Хольта та Хольта-Вінтерса. Якщо сезонність та тенденція явно відсутні у часовому ряді, то можна без сумнівів використовувати спосіб простого експоненційного згладжування. Якщо ви

спостерігаєте лінійну тенденцію в ряді, варто використовувати метод подвійного експоненційного згладжування або метод Хольта, якщо тенденція нелінійна - допомагатиме метод потрійного згладжування. Якщо ж в працюєте з рядом, в якому присутні сезонність і тенденція разом, при чому незалежно в який спосіб вони комбінуються (мультиплікативно чи адитивно) варто користуватись методом Хольта-Вінтерса. В фінансово-економічній сфері саме такі типи рядів можна зустріти найчастіше.

Розглянемо ближче ідею розглянутих методів. Метод простого експоненційного згладжування передбачає лише згладжування часового ряду ковзною середньою з експоненційними вагами. Зрозуміло, що назва методу говорить про те, що дані згладжуються за допомогою зваженої середньої, для якої ваги змінюються за експоненційним законом. Ковзна середня здатна краще характеризувати значення процесу на кінці ніж на початку інтервалу дослідження. Не потребує пояснень факт, що найновіші значення часового ряду мають найбільший вплив на результат прогнозу, саме тому їм прийнято давати більшу вагу порівняно з початковими. Хоча початкові дані також не можна цілком виключити з аналізу так як вони несуть собою певну інформацію про досліджуваний економічний процес. Формалізуємо метод простого експоненційного середнього: допустимо що у нас є ряд спостережень зі змінною y : y_1, y_2, \dots, y_n . Тоді просте експоненційне середнє можна порахувати за формулою:

$$S_n^1 = \sum_{j=0}^{n-1} \alpha(1-\alpha)^j y_{n-j} \quad (2.1)$$

де S_n^1 – експоненційне середнє першого порядку для n – го періоду; α – коефіцієнт згладжування: $0 < \alpha < 1$, $\alpha(1-\alpha)^j \geq 0$, $\sum_{j=0}^{n-1} \alpha(1-\alpha)^j \Rightarrow 1$.

Просте експоненційне середнє є прогнозом досліджуваного значення в періоді $(n+1)$, тобто:

$$\hat{y}_{n+1} = S_n^1. \quad (2.2)$$

Також для простого експоненційного середнього дійсна така формула рекурсії:

$$S_n^1 = \alpha y_n + (1 - \alpha) S_{n-1}^1. \quad (2.3)$$

Бачимо, що для використання формули нам необхідно визначити початкове значення. Це і є один з найбільших недоліків методу - відсутність формалізації при виборі значення параметру згладжування α . Далі розглянемо деякі найпоширеніші підходи до вибору параметра згладжування.

2.3.1 Методи вибору α

Безпосередньо від значення параметру α залежить наскільки якісним буде результат прогнозу. Чим ближче значення α до одиниці, тим більший вплив на прогноз матимуть останні спостереження. Якщо ж значення α близьке до нуля, то більший вплив матимуть старіші значення ряду, так як тоді ваги спадають повільніше. Поширеними значеннями параметра є межі α від 0.1 до 0.3, така рекомендація була запропонована вченим Брауном для прогнозування практичних задач, проте але його рекомендація нічим особливо не виявилась обґрунтованою [11].

Нижче представлено приклад ітеративної процедури для вибору значення α [9]:

- спочатку обираємо критерій оцінки якості прогнозу. Це можуть бути СКП, СПП, АСПП та інші;
- вибираємо крок зміни для параметра α наприклад з кроком 0.1. Після цього визначаємо підмножину значень α , яка дорівнює: [0; 0.1; 0.2; ...; 0.9; 1];
- Підбираємо початкове наближення, за одним із методів описаних вище
- розраховуємо простим методом експоненційного згладжування кожне значення α з побудованої підмножини
- обираємо один з критеріїв оцінки якості прогнозу та обчислюємо його значення: МАП, МАПП та ін.;
- оцінюємо отриманий набір значень α , та шукаємо результат який відповідає мінімальному значенню критерію.

Звісно описаний спосіб полегшує завдання вибору параметра згладжування, але які ж недоліки описаного методу? На мою думку, їх всього кілька: необхідність збереження всього ряду спостережень замість одного попереднього значення, велика кількість операцій для обчислень, а також необхідність збору всіх прогнозних значень та помилок прогнозу для всіх значень параметру α .

2.3.2 Вибір початкового наближення

Особливу увагу варто приділити експоненційним середнім різних порядків, що відіграють важливу роль для методів простого, подвійного та

потрійного експоненційного згладжування. Їх знаходять за рекурсивною формулою, проте тут виникає проблема зі значеннями початкових наближень. На сьогоднішній день не існує визначеного єдиного способу визначення початкових наближень, часто їх вибирають в залежності від конкретної предметної області дослідження. Для прогнозування біржового курсу часто вибирають перше значення ряду, якщо прогнозується вартість товарів чи послуг, частіше використовують середнє значення досліджуваного ряду, так як ціна зазвичай коливається навколо свого середнього значення або багато інших варіантів. Цікавим способом вибору початкового наближення також є варіант взяття середнього кількох початкових або кінцевих значень ряду. Ще можна взяти наближення як параметр :

$$S_0^1(y) = \frac{\sum_{i=1}^n y_i}{n} \quad (2.4)$$

2.4 Метод групового врахування аргументів

2.4.1 Основні ідеї методу

Не менш популярним методом прогнозування сьогодні є метод групового врахування аргументів (МГВА), він ґрунтується на комбінації опорних функцій, системі правил ускладнення моделі, а також критерію селекції та методу регуляризації по зовнішнім критеріям. Спочатку проводиться вибір моделей-претендентів, далі з них відсівають моделі, що не пройшли поріг селекції. В цьому і полягає алгоритм самоорганізації. Запишемо детальніше по крокам самоорганізаційну структуру алгоритму моделі:

- обробити спостереження базуючись на системі обраних опорних функцій для зменшення кількості претендентів;
- згенерувати множину моделей-претендентів;
- обчислити критерії селекції, які є зовнішніми доповненнями, а також знайти моделі(і) оптимальної складності.

У метод групового врахування аргументів є багато різних модифікацій, вони відрізняються:

- кількістю рядів селекції;
- кількістю рівнянь;
- обрахунком залишків (наявність чи відсутність);

Також методи групового врахування аргументів можна поділити на однорядні та багаторядні за кількістю рядів селекції. Багаторядні алгоритми зазвичай використовуються при наявності недовизначеності задачі моделювання, тому що регресійні моделі не можуть бути використані - для них не можливо побудувати адекватну процесу модель що лежить за межами інтерполяції. Однорядний МГВА прийнято використовувати при перевизначених задач самоорганізації за дослідними даними. Але й багаторядні ряди при перевизначених задачах також можуть використовуватись, і навіть часто дозволяють отримати кращий результат.

2.4.2 Зовнішні критерії оптимальності

Для селекції моделей популярні критерії незміщеності та регулярності, що і використовуються в роботі, вони використовуються як зовнішні доповнення. Загальним для сформульованих нижче критеріїв є те, що вони визначаються базуючись на перевірочній вибірці, а також дають можливість

повторення отриманого результату (тобто відповідають вимогам відтворюваності).

Вся вибірка для реалізації N ділиться на реалізацію навчальної вибірки N_A , за допомогою якої модель навчається на потім оцінюються параметри моделі, та також реалізацію перевіркової вибірки N_B , за допомогою якої відбувається перевірка якості моделі та вибір найкращої моделі.

За допомогою критерію регулярності визначається середньоквадратичне відхилення моделі на перевіірочній вибірці:

$$\Delta^2(B) = \frac{\sum_{t \in N_B} (y_t^M - y_t)^2}{N_B} \rightarrow \min. \quad (2.5)$$

Якщо розглядати те, що при постійному комплексі умов хороша апроксимація в минулому гарантує хорошу апроксимацію для найближчого майбутнього, тоді критерій регулярності можна використовувати для короткострокового прогнозу. Варто відмітити, що розв'язок, отриманий на більш нових реалізаціях, дає лише невелике відхилення від вхідних даних. Хоча при цьому в процесі селекції можуть бути загубленими важливі змінні, але їх вплив непрямо буде врахований через інші змінні.

Розглянемо детальніше критерій незміщеності (або як його називають, несуперечності) моделі. Основною ідеєю є те, що для одного об'єкту дослідження по різних вибіркам даних, отриманих від нього за інших рівних умов, мають бути отримані близькі моделі, за допомогою яких можна визначити поведінку об'єкту.

Критерій записуються наступним чином:

$$n_{cm} = \frac{1}{R_1 + R_2} \cdot \sum_{r=1}^{R_1 + R_2} (z_r^* - z_r^{**})^2 \quad (2.6)$$

де R_1, R_2 - розміри першої та другої підвибірок даних відповідно; z_r^*, z_r^{**} — значення прогнозу першої та другої моделі відповідно до всіх точок вибірок.

В моїй роботі для відсіву моделей на ітерації селекції використовувався критерій, що являє собою опуклу комбінацію критерію несуперечності та регулярності у наступному вигляді:

$$K_{\Sigma} = \alpha \cdot \Delta^2(B) + (1 - \alpha) \cdot n_{cm} \quad (2.7)$$

де α - ваговий коефіцієнт, $0 \leq \alpha \leq 1$.

2.4.3 Опис алгоритмів селекції

Далі розглянемо поліноміальні алгоритми методу групового урахування аргументу. припустимо, що повний опис об'єкту задається наступною залежністю

$$y = \varphi(x_1, x_2, \dots, x_n) \quad (2.8)$$

Замінімо цей вираз базисом часткових поліноміальних описів:

$$y_1 = f(x_1, x_2), y_2 = f(x_1, x_3), \dots, y_s = f(x_{n-1}, x_n), s = C_n^2. \quad (2.9)$$

Різноманітні алгоритми МГВА розрізняються за видом функції часткового опису $f(\cdot)$. В роботі розглядаються три види часткових описів:

- лінійний частковий опис

- неповний квадратичний частковий опис:

$$f(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j \quad (2.10)$$

- квадратичний частковий опис:

$$f(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (2.11)$$

де a_i – дійсне число.

В комбінаторних (однорядних) алгоритмах виконується перебір всеможливих моделей із заданого базису із вибором кращої з цих моделей за заданим критерієм селекції (або за їх комбінацією). При переборі складність часткових моделей, тобто число аргументів, поступово нарощується від 1 до максимального числа n (числа аргументів базисного набору функцій). Таким чином, загальна схема комбінаторного алгоритму включає в себе наступні операції:

- за МНК визначаються коефіцієнти всіх часткових моделей при складності $s = 1, s = 2, \dots, s = n$;
- для кожної з них обраховуються значення зовнішнього індивідуального або комбінованого критерія селекції;
- єдина модель оптимальної складності вибирається з умови мінімуму критерія.

Можна сказати, що комбінаторний алгоритм МГВА заснований на методі повної математичної індукції, оскільки при цьому не пропускається жоден із можливих варіантів моделі, закладених в початковому базисі.

Розглянемо, як ця загальна схема реалізується в поліноміальних алгоритмах. Тут опорним набором аргументів виступає n членів деякого поліному заданого максимального степеня від заданого числа змінних.

Наприклад, для випадку трьох змінних повний квадратичний поліном має вигляд:

$$q = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_1^2 + a_5 x_2^2 + a_6 x_3^2 + a_7 x_1 x_2 + a_8 x_1 x_3 + a_9 x_2 x_3. \quad (2.12)$$

Поліноми такого типу є лінійними відносно коефіцієнтів, для визначення яких застосовується МНК.

Відповідно до наведеної схеми алгоритм працює так. Спочатку визначаються всі моделі при $s = 1$, тобто ті, що складаються з одного аргумента:

$$q_1 = a_0, q_2 = a_1 x_1, q_3 = a_2 x_2, \dots, q_{10} = a_9 x_2 x_3. \quad (2.13)$$

Кількість таких моделей – C_{10}^1 . Після цього розглядаються всі можливі моделі при $s = 2$, що містять два аргументи:

$$\begin{aligned} q_{11} &= a_0 + a_1 x_1, & q_{12} &= a_0 + a_2 x_2, \dots, \\ q_i &= a_0 + a_9 x_2 x_3, \dots, & q_j &= a_1 x_1 + a_2 x_2, \dots, \\ q_k &= a_1 x_1 + a_9 x_2 x_3, \dots, & q_{45} &= a_8 x_1 x_3 + a_9 x_2 x_3. \end{aligned} \quad (2.14)$$

Всього параметрів буде $C_{10}^2 = 45$. Аналогічно будуються решта моделей. Таким чином, загальна кількість p_n всеможливих часткових моделей, побудованих із n аргументів повного полінома методом повного комбінаторного перебору дорівнює

$$p_n = \sum_{s=1}^n C_n^s = 2^n - 1. \quad (2.15)$$

Очевидно, що при збільшенні n кількість моделей зростає дуже швидко – додавання одного аргументу збільшує час розрахунків удвічі.

Структурна схема багаторядних алгоритмів наведена на рисунку 2.2.

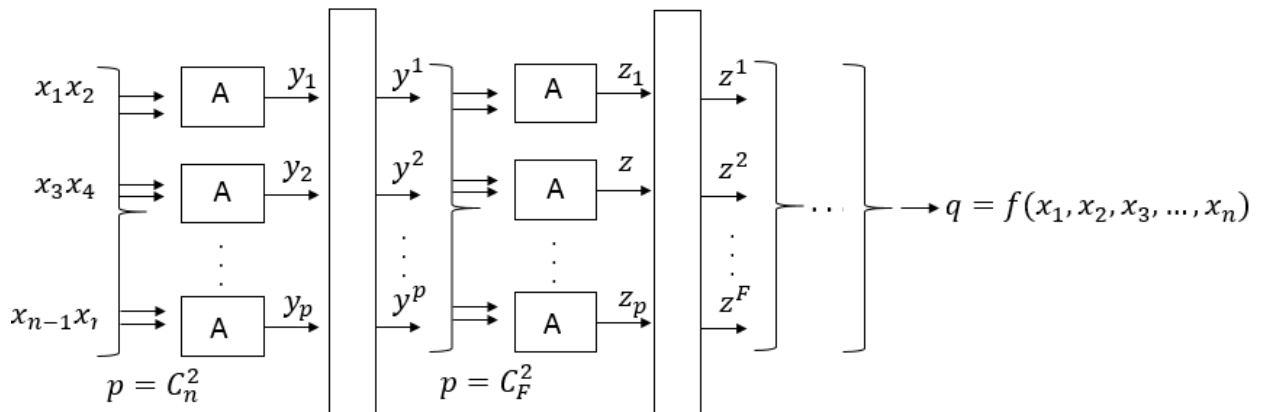


Рисунок 2.2 – Схема багаторядного МГВА

Вважається, що початковий склад аргументів, з якого починається процедура багаторядної селекції моделі процесу, будується на так званому нульовому ряді алгоритму. Наприклад, в класі алгебраїчних функцій найбільш загальною моделлю є поліном Колмагорова-Габора від k змінних:

$$q = a_0 + \sum_{i=1}^k a_i z_i + \sum_{i=1}^k \sum_{j=1}^k a_{ij} z_i z_j + \sum_{i=1}^k \sum_{j=1}^k \sum_{l=1}^k a_{ijl} z_i z_j z_l + \dots \quad (2.16)$$

Після перетворення в ньому всіх доданків, отримаємо лінійний поліном

$$q = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n. \quad (2.17)$$

Члени цього поліному i є аргументами, побудованими на нульовому ряді алгоритму.

Далі покажемо як працює багаторядний алгоритм. На першому ряді на основі даних таблиці спостережень будуються часткові описи від всіх

попарних комбінацій початкових аргументів, що наближують за МНК вихідну змінну q

$$y_1 = f_1(x_1, x_2), y_2 = f_2(x_1, x_3), \dots, y_k = f_k(x_{n-1}, x_n), \quad (2.18)$$

із яких за зовнішнім критерієм відбирається F_1 кращих за зовнішнім критерієм;

На другому ряді отримані змінні приймаються в якості вхідних аргументів і знову будуються всі часткові моделі від двох аргументів

$$z_1 = \phi_1(y_1, y_2), z_2 = \phi_2(y_1, y_3), \dots, z_l = \phi_l(z_{F_1-1}, z_{F_1}), l = C_{F_1}^2. \quad (2.19)$$

Із них за критерієм обирається F_2 кращих моделей, що використовуються в якості аргументів для наступного ряду селекції. Цю процедуру повторюють пока є зменшення значення критерію. Саме так в багаторядних алгоритмах МГВА відтворюють схему масової селекції. В них наявний генератор описів, що ускладнюються із ряду в ряд ("схрещування") та порогові відбори кращих з них ('селекція'), тобто, подібно до генетичних алгоритмів, реалізується гіпотеза селекції [11].

2.5 Метод подібних траєкторій

Для 1-вимірної часової послідовності з малим рівнем шуму розглянемо метод непараметричного прогнозування. Зазвичай на практиці приймають, що послідовність генерується нелінійною динамічною системою заданою наступними рівняннями:

$$z_{t+1} = h(z_t, u_t), \quad (2.29)$$

$$y_t = g(z_t) \quad (2.30)$$

де $h(z_t, u_t)$ та $g(z_t)$ - нелінійні незалежні від часу функції, які мають обмежені похідні; $u_t \in R^{d_s}$ - визначена функція змінних t ; $y_t \in R^1$ та $z_t \in R^{d_s}$. Вважається, що $h(z_t, u_t)$, $g(z_t)$, u_t та d_s - невідомі, а прогноз має бути побудований виключно на згенерованій системою часовій послідовності $[y_1, y_2, \dots, y_n]$.

Майже у всіх нелінійних моделей, незалежно є вони локальними чи глобальними, є параметри які мають бути визначені перед тим як будувати модель чи використовувати її. Працюючи з системами типу “black box”, про процес досліднику невідомо нічого, що могло б допомогти дізнатись ці параметри. Якщо можливо, потрібно обирати параметри з умови максимізації деякої міри точності моделі. Незважаючи на те, що порахувати параметри на комп’ютері сьогодні можна дешево і швидко, максимізувати модель глобально по всіх параметрам майже у всіх випадках і досі є майже недосяжно. Альтернатива яка приходить на думку - нелінійна оптимізація - в загальному випадку неприпустима (чи недосяжна) через дуже велику кількість локальних мінімумів.

Існують методи які в деяких випадках можуть застосовувати для знаходження допустимих значень деяких з цих параметрів. Також існуюча теорія пропонує непогані оцінки або, більш частіше, обмеження на значення параметрів. На жаль, ці варіанти в більшості випадків не можна застосувати до всіх параметрів моделі.

Досить часто точність моделі дуже чутлива до одних параметрів і менш чутлива до значень інших параметрів. Досвідчений дослідник здатен одразу вибрати прийнятні значення цим параметрам. На жаль, цей процес вибору значень параметрів є суб’єктивним і не допускає узагальнення. А значить точність моделі може сильно залежати від вміння та досвідченості людини, що впливає на вибір. Також популярно для декількох значень параметрів моделі

використовувати грубу оптимізацію аніж вимірювати точність моделі. Звісно використання такої оптимізації обмежується обчислювальними ресурсами, і, як результат, страждає точність моделі, тут потрібно звертати увагу на потреби - чи найбільше значення має точність наступного кроку, наприклад, чи загальна точність якомога більшої кількості кроків.

Також потрібно обрати підходящої міри схожості, тобто метрику, що є дуже важливим рішенням, хоча й не є складним. Адже є лише одна найрозповсюдженіша метрика – квадрат евклідової відстані між вхідним вектором x_t та цільовим вектором q :

$$d_E(q_t, x_t) = (q_t - x_t)^T (q_t - x_t) \quad (2.31)$$

В цьому випадку ми маємо один параметр – m , розмірність. Хоча вибір більш загального вигляду метрики з більшою кількістю параметрів може надмірно вплинути на точність моделі. Наприклад, зважена евклідова відстань

$$d_\Lambda(q_t, x_t) = (q_t - x_t)^T \Lambda (q_t - x_t) \quad (2.32)$$

може використовуватись як метрика, якщо $\Lambda \in R^{m \times n}$ - додатна напіввизначена матриця. Однак, короткі часові послідовності часто бувають занадто малими для оцінки m^2 параметрів. Якщо розглядати з іншого боку, для великих вибірок не вигідно з обчислювальної точки зору оптимізувати таку кількість параметрів. Метрика d_Λ визначається параметрами λ та m . При досить великих m та досить малих λ вплив найбільш віддалених компонентів x_t буде незначним. Тому, якщо модель будується при досить великому значенні m , тоді на точність моделі впливає лише λ .

Оскільки точність розгляданої моделі залежить від λ , зростає вартість обчислення при знаходженні найкращого значення λ . Це і вважається критичним моментом запропонованого методу. Якщо розглядати усі вищеописані методи, побачимо що більшість нелінійних моделей в загальному випадку потребує дуже багато обчислень для знаходження глобального екстремуму за структурними параметрами (наприклад пошук кількості нейронів для схованого шару нейромережі). Чим ж відрізняється метод подібних траєкторій? Для нього запропоновану модель можна побудувати та ефективно виконати глобальну оптимізацію за параметрами моделі, що робить вибір параметрів ключовим моментом.

До переваг методу можна віднести те, що завжди можна вибрати прийнятне значення розміру вікна при візуальному дослідженні часової послідовності. За відсутності шуму, типова одновимірна часова послідовність, буде містити періодичні коливання приблизно однакового періоду. Розмір вікна w потрібно вибрати таким чином, щоб воно охоплювало декілька коливань. Більш велике вікно потребує більших обчислень при невеликому зростанні точності. При малій довжині часової послідовності велике вікно значно скорочує кількість точок, що можуть бути використані для побудови моделі:

$$n_c = n - m\tau . \quad (2.33)$$

І навпаки, занадто мале вікно зменшить точність моделі.

Нехай ми визначились з розміром вікна w . Треба вибрати значення часу дискретизації τ . Єдиним виправданням вибору найменшого значення τ є точність d_λ .

Для фіксованого значення w , вибір τ впливає на точність зваженої квадратичної похибка апроксимації ISE_λ . Малі значення збільшують точність ISE_λ , однак точність моделі не обов'язково збільшиться. З іншого боку, зменшення τ означає збільшення m , а отже зростання вартості обчислення. Беручи до уваги це протиріччя, τ треба обирати якомога меншим з обмеженням по обчислювальних можливостях.

Якщо підсумувати вищесказане, то за обраним критерієм вибирається ділянка траєкторії “найближчу” до ділянки, що передуює прогнозній точці. Потім робиться прогноз за формулою:

$$\hat{y}(k+s)=y(I+J+s-1), \quad (2.34)$$

де

$$I = \min_j \left\{ \sum_{i=1}^p |y(j+i-1) - y(n-p+i)| \right\} \quad j = 1, 2, \dots, n-p;$$

$$J = \min_i |y(I+i-1) - y(n)| \quad i = I, I+1, \dots, I+p-1.$$

Формально можна записати наступним чином:

$$Y_1 = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} Y_2 = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_{p+1} \end{pmatrix} \dots Y_k = \begin{pmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+p-1} \end{pmatrix} \dots Y_y = \begin{pmatrix} y_{n-p+1} \\ y_{n-p+2} \\ \vdots \\ y_n \end{pmatrix}. \quad (2.35)$$

Знаходимо найближчу точку із умови мінімальної відстані:

$$Y_k = \arg \min_j d(Y_n, Y_j). \quad (2.36)$$

2.6. СППР

2.6.1 Актуальність створення СППР

В сучасному світі, де все змінюється з неймовірною швидкістю неможливо визначити стандарти чи методи управління які підійдуть всім. Проте, звісно, чим більшою є обґрунтованість того чи іншого методу управління для конкретного випадку, тим меншим є комерційний ризик. Ефективно вирішувати задачі планування чи контролю без хорошого прогнозування ключових показників та процесів просто неможливо. Якщо при розробці рішення про управління недостатньо обдумане прогнозування стратегії, впливу зовнішніх факторів, змін частин внутрішнього середовища об'єкта, тоді ймовірніше таке рішення буде неякісним та призведе до небажаних результатів.

Покажемо, що процес прийняття рішень вимагає створення та використання ієрархічної системи прогнозування. Зрозуміло, що рух вниз по управлінській ієрархії зазвичай супроводжується значними змінами характеру інформації та рівню її деталізації. Так, на вершині піраміди необхідно приймати неструктуровані та погано структуровані рішення при неповній інформації (або її відсутності). Чим нижчим є рівень управління, тим більше програмованими (структурованими, упорядкованими, прозорими) є рішення, для яких існує інформація у все повнішому об'ємі. Сьогодні при прийнятті будь-якого рішення ми зважуємо та аналізуємо інформацію, оскільки для будь-яких комерційних(та насправді багатьох інших!) рішень кількість інформації дуже велика, без автоматизованої системи підтримки прийняття рішень просто не обійтись - лише так можна автоматизувати більшість процесів управління, включаючи планування, контроль та аналіз ризиків. При цьому значно зростає докладність описання ситуацій та задач через можливі об'єми обробки та аналізу інформації для прийняття рішення, а відповідно зростає і ступінь автоматизації процесу прийняття рішень.

Отже, розуміємо, що прогнозуюча(чи аналізуюча) система також повинна бути побудована за ієрархічною структурою. Це дає можливість упорядкувати задачі прогнозування, створити множину необхідних для прогнозування спеціалізованих моделей, ситуацій та методів прийняття рішень на кожному рівні, а також наблизити комп'ютеризований процес прийняття рішень до того, що є звичним для нас. При цьому безперечно такі рішення будуть значно точнішими, хоча б враховуючи яка кількість параметрів чи елементів може досліджуватись та прогнозуватись одночасно [14].

2.6.2 Опис СППР

Основною метою створюється даної СППР є вибір кращого методу прогнозування для процесу, що описується часовим рядом. Очевидно, що для кожного процесу часовий ряд має свою структуру, метод побудови, статистичні параметри тощо. Тому неможливо чітко визначити параметри кожного з методів, що мають бути закладені в систему. А тому окрім вибору методу прогнозування необхідно також обрати параметри, при яких модель процесу буде достатньо адекватною для прогнозування й надаватиме точний прогноз. З надходженням нової інформації про процес (нових елементів часового ряду) характер процесу може змінитись, що може також призвести до вибору інших параметрів моделі, або, навіть, іншого методу прогнозування. Тому задача має вирішуватись циклічно, постійно аналізуючи нові дані, значить СППР матиме замкнену циклічну структуру.

Вихідною інформацією для процесу очевидно буде сам часовий ряд, кількість його елементів, характер часового ряду (стаціонарність, мінімальне та максимальне значення, перший та другий моменти). Параметрами процесу

мають бути перелік статистик, за якими буде оцінюватись якість прогнозу (критерії вибору рішень), методи, що можливо застосувати до прогнозування, методи попередньої обробки даних тощо.

Проаналізувавши сутність процесу прогнозування розглянемо частини процесу, які будуть виконуватись за допомогою майбутньої СППР. Природньо буде доручити ЕОМ етапи, які потребують значних об'ємів обчислень. Однак СППР може взяти на себе також частину рішень, що має приймати ОПР. Для рішень, що будуть доручені СППР мають бути розроблені чіткі та детальні алгоритми. Але варто пам'ятати що такий підхід має свої плюси та мінуси.

Безумовно, остаточне рішення приймає людина при наявності усіх результатів аналізу та прогнозування. СППР має видати ОПР рекомендацію, проте як ОПР приймає рішення, керуючись цією рекомендацією чи взагалі не беручи її до уваги. Для проектованої системи обчислювальними процедурами будуть:

- методи попередньої обробки даних;
- методи побудови прогнозів;
- обчислення критеріїв якості прогнозу;
- побудова графічних представлень вихідних даних та результатів.

Всі ці процедури мають виконуватись за рахунок обчислювальної системи (програми). Крім того, слід передбачити ситуації, в яких система має видати користувачу (ОПР) рекомендації. Такими ситуаціями можуть бути:

- вибір процедури попередньої обробки даних;
- попередній відбір методів на основі можливостей їх застосування;
- визначення діапазону параметрів для кожного з методів;
- вибір кращого прогнозу серед результатів прогнозування конкретним методом;
- вибір кращого прогнозу часового ряду серед кращих прогнозів всіх застосованих методів.

Повний опис СППР бачимо нижче в таблиці 2.1.

Таблиця 2.1 – Опис СППР

Ситуація з прийняття рішення:	прогнозування часового ряду різними методами	
Динаміка задачі:	ітерації в замкнутому циклі	
Ситуативні цілі:	обрати кращий прогноз по кожному з методів прогнозування, обрати кращий метод прогнозування для даного часового ряду	
Критерії оцінки:	порівняння результатів прогнозу за допомогою статистичних критеріїв (СКП, САПП, МАПП, МіАПП)	
Фізичний процес:	будь-який процес, що описується часовим рядом	
Інформаційна підтримка:	-	
Входи:	Виходи:	Параметри:
часовий ряд, довжина ряду, статистичні параметри ряду, кількість кроків прогнозу	прогнозні значення, параметри кращих прогнозів для кожного методу, значення критеріїв якості для кожного прогнозу	методи попередньої обробки даних, методи прогнозування, критерії вибору кращого прогнозу
Проміжний аналіз:	Вибір процедур попередньої обробки даних, попередній відбір методів прогнозування, визначення діапазонів перебору параметрів кожного методу, визначення кращого прогнозу кожного методу, визначення кращого прогнозу часового ряду.	
Представлення ситуації:	ОПР приймає рішення виходячи з характеру часового ряду (представленого рядом чисел та графічно), отриманих значень критеріїв прогнозу (у вигляді таблиці) та характеру прогнозу (представленого рядом чисел та графічно)	

Продовження таблиці 2.1

Необхідні судження:	якщо ряд потребує попередньої обробки, обрати методи обробки; обрати методи, які можливо застосувати до прогнозування даного ряду; обрати можливі діапазони параметрів для кожного з обраних методів; порівнюючи отримані значення статистичних критеріїв обрати кращий метод прогнозування
---------------------	---

Отже, система призначена для виконання прогнозів часових рядів одним методом чи набором методів та порівняння результатів прогнозування різними методами. Надалі можливо покращення результатів за рахунок комбінування оцінок прогнозів отриманих різними методами.

2.6.3 Архітектура СППР

Для того, щоб правильно розв'язати задачу проектування СППР, необхідно максимально структурувати описання процесу прийняття рішення. Основною проблемою при проектуванні СППР є аналіз та виявлення процесу прийняття рішення ОПР, визначення обмежень, що накладаються на процес прийняття рішення, а також вибір методів та обчислювальних засобів, які дозволять зняти подібні обмеження. В загальному випадку проектування СППР складається з трьох наступних етапів:

- а) Декомпозиція процесу прийняття рішення на елементарні операції та описання виконання цього процесу особою, що приймає рішення.
- б) Аналіз конкретної задачі стосовно прийняття рішення та проектування СППР на функціональному рівні.
- в) Докладна специфікація функцій системи, її реалізація та верифікація.

Функціональне проектування СППР представляє собою когнітивний інженерний процес, який включає в себе двохетапне узгодження даних, що надходять від конкретного описання ситуації з прийняття рішень, з технологічною базою системи підтримки прийняття рішень. В більшості випадків прийняття рішень ОПР зацікавлена в порівнянні наслідків прийняття альтернативних рішень. Порівняння альтернатив зв'язане, як правило, з складними кількісними розрахунками, необхідністю використання спеціальних знань та специфічних моделей для оцінювання альтернатив. Інформація поступає в СППР не менше ніж у двох видах: дані, що характеризують значення конкретних атрибутів або факти стосовно реальної ситуації, та знання, які структурно та семантично описують попередній досвід та дають можливість екстраполювати нові ситуації. Тому проектування СППР – комплексна задача, при розв'язку якої необхідно враховувати всі можливі варіанти реалізації для вибору кращої архітектури та стратегії роботи системи.

Процес проектування починається з того, що вибирається задача, для розв'язування якої необхідно створити систему підтримки прийняття рішень або ця задача ставиться ОПР, яка вже знайома з проблемою прийняття конкретних рішень з власного досвіду. На першому етапі виконується декомпозиція задачі елементарні операції та описання виконання цього процесу особою, що приймає рішення. Основна мета цього етапу полягає в наступному:

- визначити перепони («вузькі місця»), які необхідно подолати при прийнятті рішень за допомогою СППР, що проектується;
- визначити набір комп'ютерних алгоритмів, які необхідно використати для подолання перепон, пов'язаних з оперативним прийняттям правильних рішень.

Для того, щоб правильно розв'язати задачу проектування СППР, необхідно максимально структурувати описання процесу прийняття рішення. Така структура представляється докладним, але чітко визначеним протоколом, в якому вказується, які дані необхідно зібрати, та описати всі часткові рішення,

що повинні бути прийняті при проектуванні СППР. Результатом виконання цього етапу є структурована таблиця, в яку зводяться всі результати стосовно декомпозиції проблеми проектування СППР.

Три наступних етапи представляють собою фазу аналізу та функціонального проектування СППР. Спочатку необхідно проаналізувати саму задачу, по відношенню до якої необхідно приймати рішення, та підхід ОПР до її розв'язку. В процесі цього аналізу визначаються фактори, що створюють перепони до розв'язку задачі, якщо це робиться без комп'ютера. Простим інструментом для виконання цієї задачі є написання чіткого переліку загальної інформації, яка необхідна для підтримки прийняття рішення. Цей перелік можна порівняти з таблицею вимог, що була створена на першому етапі. Табулювання ситуацій, пов'язаних з прийняттям рішень, дозволяє конкретизувати пов'язані з ними конкретні обмеження. Таким чином, результатом виконання другого етапу проектування СППР є перелік конкретних труднощів, що виникають при прийманні рішень стосовно розв'язку конкретної проблеми.

Функціональна схема розроблюваної СППР представлена на рисунку 2.3 нижче.

Розроблювана СППР представляє собою набір структурних підсистем: мовна система (МС), процесор обробки запитів та даних, які надходять від МС, далі процесор взаємодіє з базою даних (БД) та системою обробки даних та генерування результатів (СОДГР), яка далі співпрацює з пам'яттю для тимчасового зберігання результатів. Далі пам'ять видає інформацію до системи представлення результатів. Мовна система забезпечує зв'язок між користувачем і всіма компонентами комп'ютерної системи. За її допомогою користувач формулює проблему і керує процесом її рішення, використовуючи запропоновані мовною системою синтаксичні та семантичні засоби.

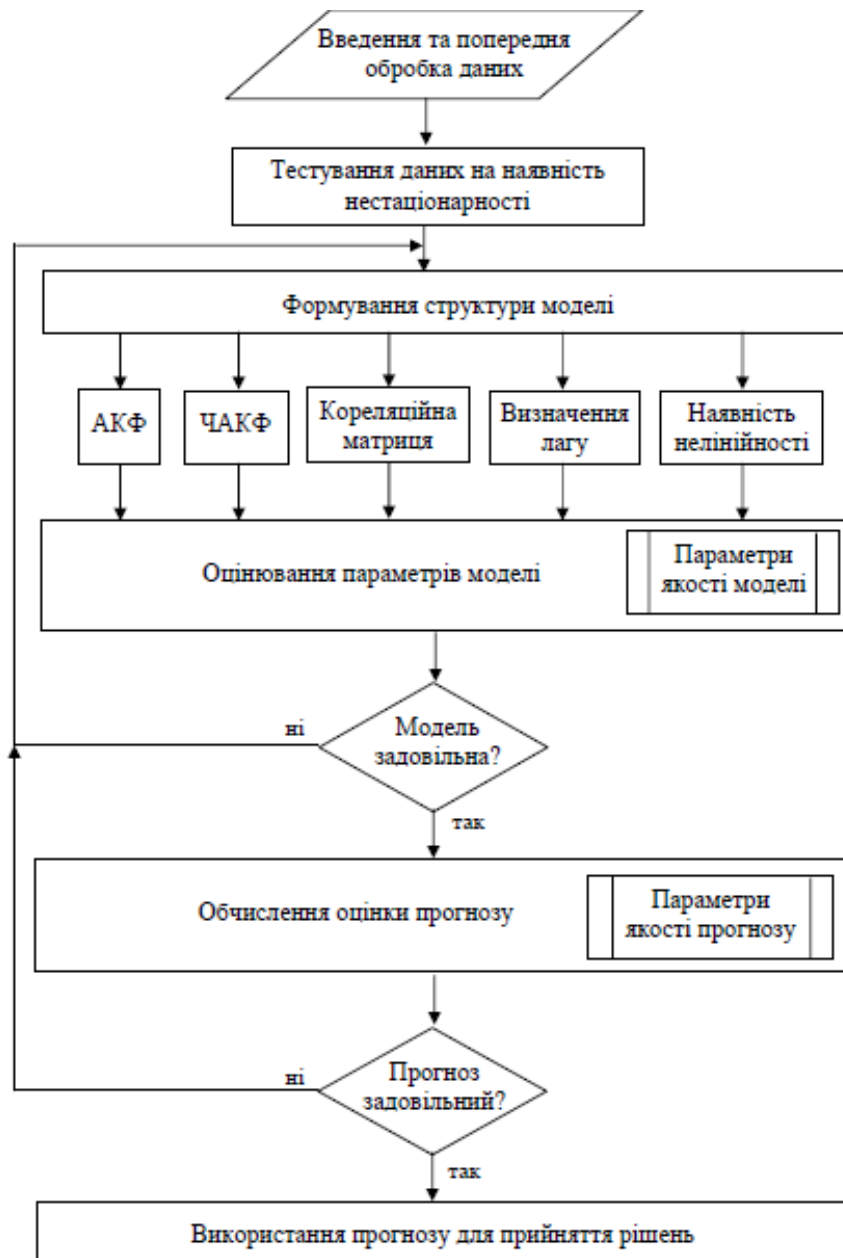


Рисунок 2.3 – Функціональна схема СППР

Головною системою СППР є розробка системи обробки даних та генерування результатів (СОДГР) з точки зору її основного призначення – підтримки прийняття рішень при розв’язанні задач моделювання та прогнозування. Розглянемо її функції, яка приймає коректні запити від мовної системи і виконує дії, які задаються запитом:

- поповнює (при необхідності) базу даних і знань;
- вибирає алгоритм (алгоритми) обробки даних з бази даних і знань;

- застосовує алгоритм МНК до даних, що містяться в базі даних, з метою побудови математичної моделі та обчислення прогнозу;
- використовує критерії адекватності моделі (які також містяться у базі даних) з метою визначення ступеня адекватності побудованої моделі;
- застосовує критерії якості прогнозу для визначення кращого прогнозованого значення;
- передає отримані результати обчислень в систему представлення результатів, яка представляє отриманий результат у формі, заданій користувачем;
- зберігає, при необхідності, частину отриманих результатів в короткостроковій пам'яті для подальшого використання в поточній сесії використання СППР.

Запропонована СППР призначена для прогнозування, саме тому в процесі підтримки прийняття рішень будуть використовуватись такі обчислювальні алгоритми:

- алгоритми попередньої обробки даних (логарифмування, нормування, фільтрація, заповнення пропусків, обробка великих імпульсних значень);
- оцінювання параметрів (коефіцієнтів) математичних та/або статистичних моделей;
- алгоритми обчислення статистичних параметрів, які характеризують якість (адекватність) побудованої моделі;
- алгоритми обчислення оцінок прогнозованих значень;
- алгоритми обчислення статистичних показників якості прогнозів;
- алгоритми вибору кращої моделі та кращого прогнозу на основі розрахованих показників адекватності моделі та якості прогнозу, використовується комплексний метод, що враховує одразу оцінки якості моделі та якості прогнозу.

Бази даних СППР включають як кількісну так і якісну інформацію, що надходить від процесору обробки запитів і даних.

База моделей забезпечує гнучкість моделювання, зокрема, за рахунок використання готових блоків моделей і підпрограм. Управління моделями дає такі можливості: каталогізувати та обслужити широкий спектр моделей, які підтримують всі рівні управління; легко і швидко створювати нові моделі; пов'язувати моделі з відповідними базами даних.

2.7 Висновки до розділу 2

Зроблено опис наступних методів прогнозування часових рядів: методи експоненційного згладжування, метод групового врахування аргументів (МГВА) та метод пошуку подібних траєкторій.

Серед переваг відзначимо швидкість, простоту та математичну обґрунтованість результатів, отриманих за допомогою експоненційного згладжування; відсутність потреби задавати структуру моделі при використанні методу МГУА; та отримання моделі у вигляді математичного рівняння в результаті застосування кожного з цих методів.

Недоліком МГУА відзначена ресурсоємність цього методу, яка не дозволяє його застосувати в широкому колі задач. Метод пошуку подібних траєкторій потребує повної перебудови моделі при зміні будь-якого параметру. Однак, на довгих рядах він може надати достатньо якісний прогноз без значних ресурсоємних обчислень.

Окрім цього досліджено прогнозування за допомогою нейронних мереж, основними перевагами яких є гнучкість та адаптивність, тобто незалежність експерта від математичної моделі поведінки досліджуваного часового ряду. Недоліком є те, що модель є чорним ящиком.

Також наведено опис ймовірнісного методу прогнозування – на основі мереж Байєса. З кожним днем вони знаходять все ширше застосування в інформаційних системах обробки статистичних даних, представлених

часовими рядами і часовими перерізами, а також якісними даними, представленими експертними оцінками, лінгвістичними змінними, інтервальними значеннями і т. ін.

В наступному розділі буде виконано порівняльний аналіз результатів наведених методів при прогнозуванні деяких типових часових рядів.

РОЗДІЛ 3

ВИКОНАННЯ І АНАЛІЗ ОБЧИСЛЮВАЛЬНИХ ЕКСПЕРИМЕНТІВ

3.1 Опис даних для прогнозування

Для прогнозування та аналізу результатів використано декілька часових рядів, що представлені у вільному доступі на сайті з всеможливими економічними показниками - <http://www.investing.com>.

NYSE_Composite — біржовий індекс американських цінних паперів (акцій підприємств), є одним з найпопулярніших фондових індексів у світі. Відображає зміну вартості всіх акцій, які торгуються на Нью-Йоркській фондовій біржі. На сьогодні включає приблизно 2 000 компаній, 1 600 з яких є американськими (загальна капіталізація цих компаній становить близько 20 трильйонів доларів), а 360 представляють інші країни.

Робоча вибірка складається з 74 значень (у вихідні дні біржа не працює) індексу в період з 2 січня 2019 року по 19 квітня 2019. Період дискретизації ряду — одна доба. В роботі імітувались та досліджувались прогнози на 2 та 3 кроки. Тому, для зручності та можливості перевірки, часовий ряд був штучно зменшений і використовувався без двох (трьох) останніх значень. Саме це дозволяє нам перевірити якість прогнозу та побачити це наочно. Усі розрахунки проводились за допомогою мови статистичної обробки даних R. Графічна інтерпретація ряду наведена на рисунку 3.1.

Із графіку важко судити про стаціонарність ряду, а відповідно і про можливість застосування до нього конкретного класу моделей. Одним із способів перевірки стаціонарності процесу є тест Дікі-Фуллера, котрий належить до виду тестів на «одиничні корені».

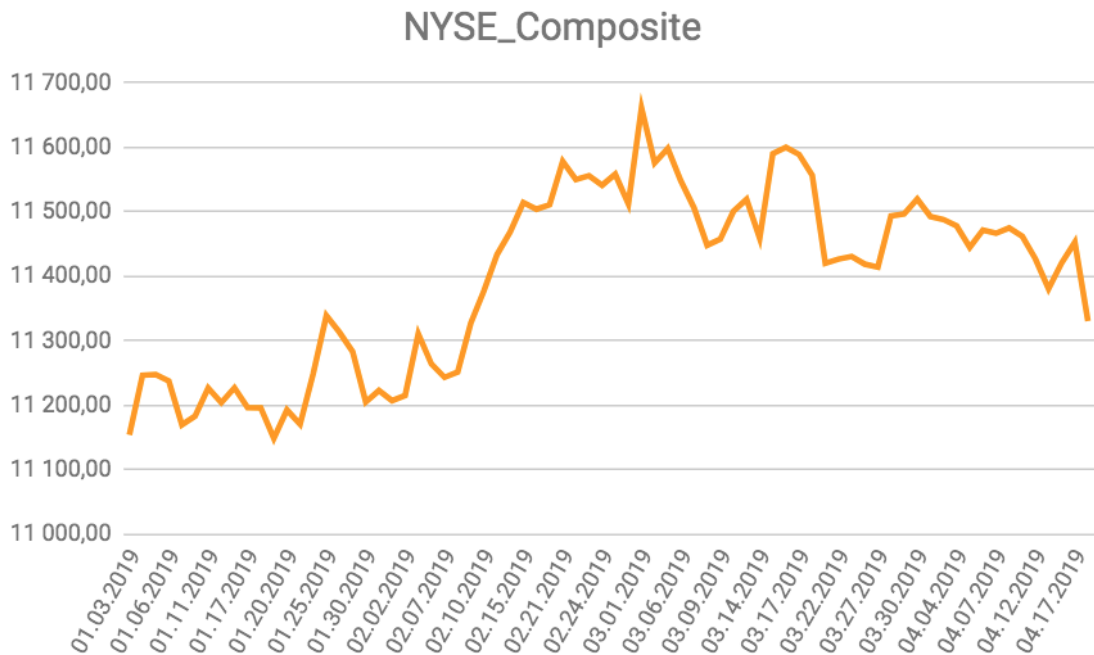


Рисунок 3.1 Графічне представлення часового ряду NYSE_Composite

Ряд має одиничний корінь (порядок інтегрованості 1), якщо його перші різниці утворюють стаціонарний ряд. Це можна записати як $y(k) \sim I(1)$. Суть тесту полягає у перевірці гіпотези H_0 про рівність одиниці коефіцієнта a рівняння 3.1

$$\Delta y(k) = (a - 1)y(k - 1) + \varepsilon(k) \quad (3.1)$$

де $\Delta y(k) = y(k) - y(k - 1)$ – оператор різниці першого порядку (різницю порядку d можна отримати послідовним застосуванням оператора першої різниці d разів). Дана статистика має розподіл Дікі-Фуллера. Якщо значення статистики лежить лівіше критичного значення при заданому рівні значущості, то гіпотеза про одиничний корінь відхиляється і процес визнається стаціонарним (в сенсі даного тесту).

В таблиці 3.1 наведені результати проведеного тесту Дікі-Фуллера для ряду NYSE_Composite .

Таблиця 3.1– Результати тесту Дікі-Фуллера для ряду
NYSE_Composite

DF-статистика	H_0	$p\text{-value}$
-1,5269	процес нестационарний	0,7685

Таким чином, імовірність зробити помилку I-го роду відхиливши гіпотезу про нестационарність становить 0,7685.

Файл EUR_USD.txt містить дані про відношення валют євро та долару. Робоча вибірка складається зі значень (у вихідні дні біржа не працює) індексу в період з 1 грудня 2018 року по 19 квітня 2019. Період дискретизації ряду – одна доба. (додаток В). Даний часовий ряд зображено на рисунку 3.2.

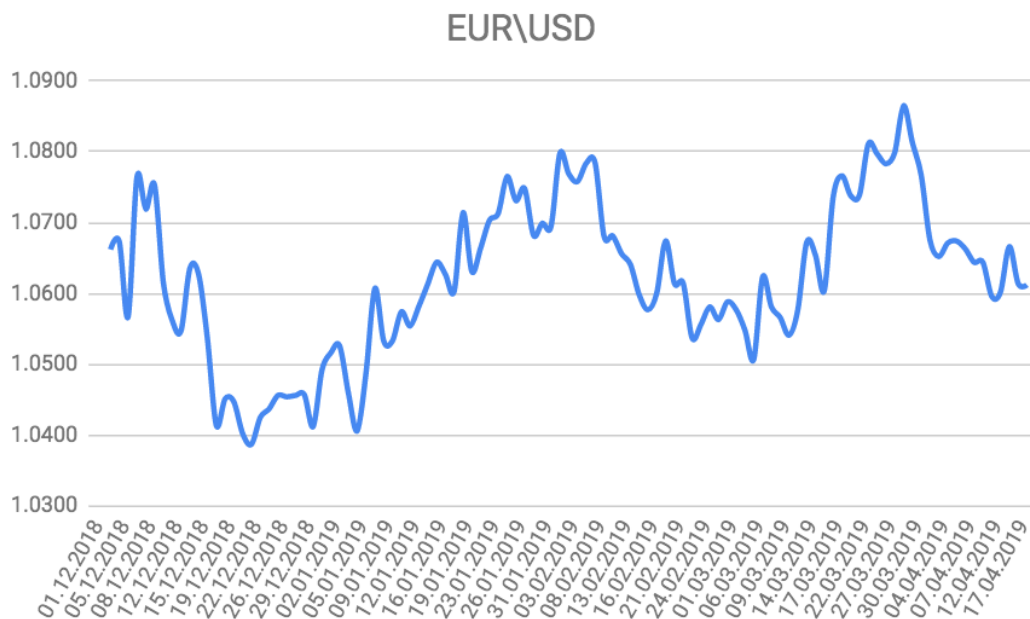


Рисунок 3.2– Графічне представлення часового
ряду відношення валют євро та долару.

Перевіримо ряд на стаціонарність за допомогою тесту Діккі-Фуллера, як і в попередньому випадку. Результати тесту наведені в таблиці 3.2.

Таблиця 3.2– Результати тесту Дікі-Фуллера для ряду EUR\USD

DF-статистика	H_0	p -value
-2,5385	процес нестационарний	0,3504

Отже, даний ряд не можна вважати стаціонарним.

Файл Wheat_prices.txt містить дані про ціни ф'ючерсів на пшеницю в США з червня 2013 року по квітень 2019. Частота дискретизації– 1 місяць. Графічна інтерпретація ряду наведена на рисунку 3.3.



Рисунок 3.3– Ціна ф'ючерсів на пшеницю в США

Файл Indexes_for_network.txt містить значення найбільш вагомих фондових індексів США з 2 січня 2017 року по 19 квітень 2019. Частота дискретизації даних – 1 день. Для побудови МБ були використані наступні:

- S&P 500 - це фондовий індекс, у кошик якого включено 500 акціонерних компаній США, що мають найбільшу капіталізацію. Список належить компанії Standard & Poor's і нею ж складається;

- промисловий індекс Доу-Джонса (англ. Dow Jones Industrial Average) - біржовий індекс цінних паперів (акцій) 30 найбільших американських підприємств. Індекс був створений для зведення інформації по

акціях індустріальних підприємств на американській біржі цінних паперів. Один з найстаріших індексів цінних паперів у США;

- NASDAQ-100 - біржовий індекс, який відображає динаміку зміни вартості акцій 100 найбільших компаній, які присутні у лістингу біржі NASDAQ. В індексі відсутні компанії фінансового сектору;

- Nasdaq Composite (досить часто трейдери називають його просто Nasdaq) - включає більше трьох тисяч компаній (як американських, так і іноземних), які входять до лістингу біржі NASDAQ. Акції будь-якої з цих компанії впливають на індекс пропорційно своїй ринковій вартості. І їх кількість значно перевищує кількість іноземних компаній, представлених в AMEX і NYSE. Через специфіку біржі NASDAQ індекс вважається важливим показником динаміки курсу паперів високотехнологічних і швидко зростаючих компаній;

- NYSE Composite — біржовий індекс американських цінних паперів (акцій підприємств), є одним з найпопулярніших фондових індексів у світі. Відображає зміну вартості всіх акцій, які торгуються на Нью-Йоркській фондовій біржі. На сьогодні включає приблизно 2 000 компаній, 1 600 з яких є американськими (загальна капіталізація цих компаній становить близько 20 трильйонів доларів), а 360 представляють інші країни.

3.2 Опис програмного продукту

Для побудови та перевірки якості прогнозів методами експоненційного згладжування, МГВА та нейронної мережі було розроблено програму на мові R— функціональній мові для статистичного моделювання та обробки даних. Даний формат реалізації був обраний завдяки наступним факторам:

- розповсюдження під GNU General public license;
- кросплатформеність (підтримуються платформи Microsoft Windows, Mac OS X, FreeBSD, Solaris та інші дистрибутиви Unix та Linux);
- зручність та швидкість роботи з даними та широкі можливості їх аналізу;
- можливість легкої інтеграції із популярними системами керування сховищами даних, наприклад, Microsoft SQL Server [24].

Для коректної роботи програмного продукту необхідно мати встановлену платформу R версії 3.1.2 та вище та апаратні можливості, що відповідають вимогам платформи R.

Продукт має модульну структуру та містить окремі скриптові файли для кожного з методів прогнозування а також для спеціалізованої візуалізації даних. Лістинг модулів програми наведений в додатку Б.

3.3 Побудова та порівняння оцінок короткострокових прогнозів

Експеримент 1. Для порівняння оцінок двокрокового прогнозу використано ряд значень індексу NYSE Composite. Для прогнозування методом групового врахування аргументів (рисунок 3.4) було використано дворядний алгоритм із розміром навчальної вибірки 45 значень (60% від загальної кількості вимірів), за даними отриманої моделі відновлено 29

значень. Для наочності результатів, числові значення перших 45 значень ряду, які використовувались для навчання алгоритму, зображені умовно – 11000.

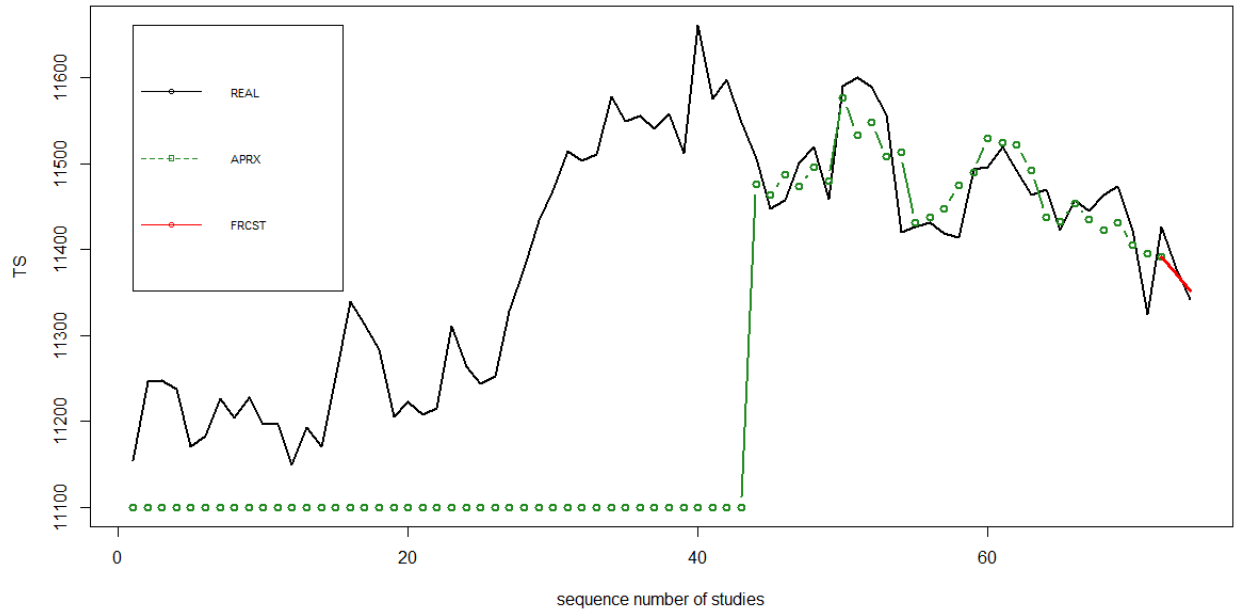


Рисунок 3.4– Двокроковий прогноз МГВА

На рисунках 3.5, 3.6 зображено результати прогнозування методом експоненційного згладжування та за допомогою нейронної мережі відповідно.

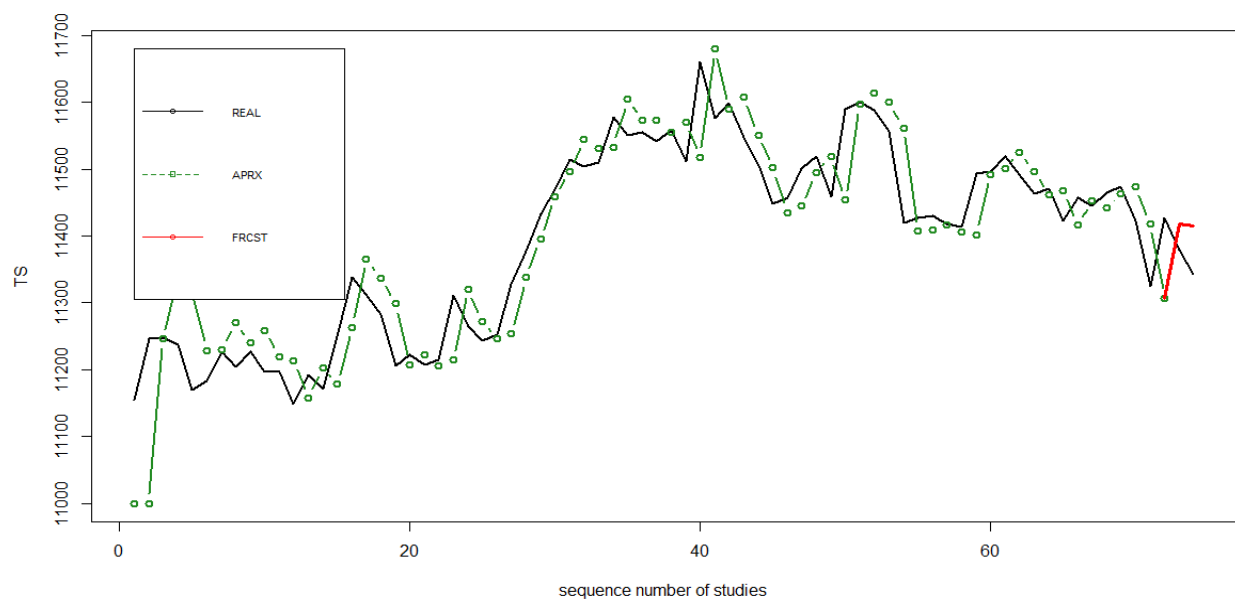


Рисунок 3.5– Двокроковий прогноз методом експоненційного згладжування

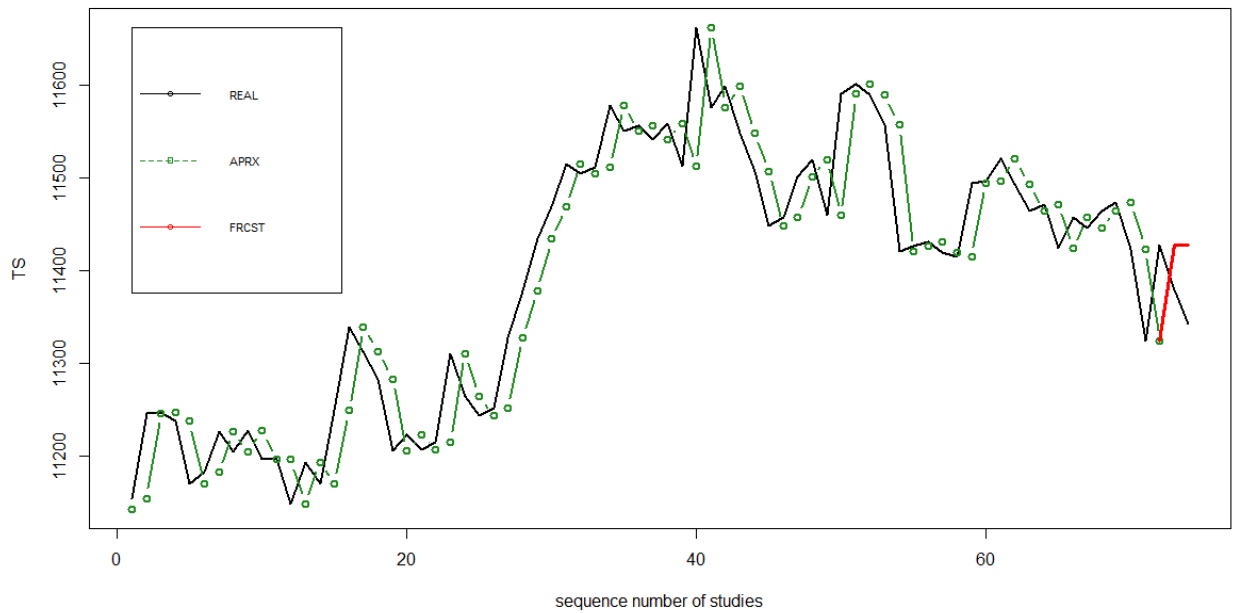


Рисунок 3.6– Двокроковий прогноз, отриманий за допомогою нейронної мережі

На відміну від попередніх методів, метод подібних траєкторій для прогнозування, не будує моделі процесу (рисунок 3.7), тому для нього відсутній графік відновлених за моделлю даних.

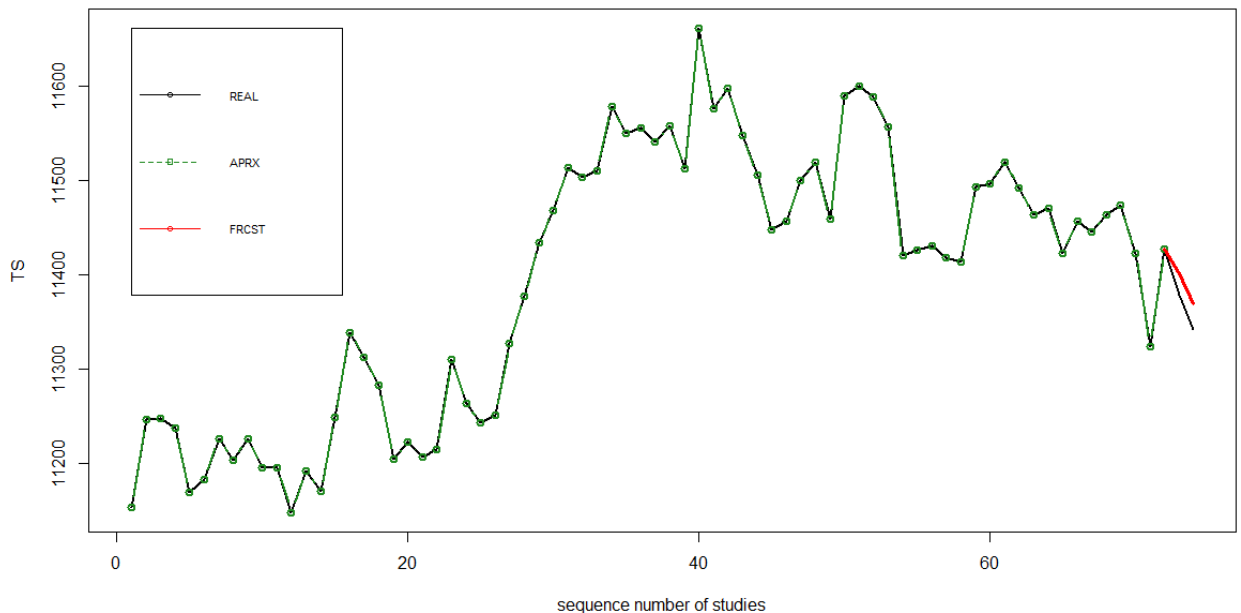


Рисунок 3.7– Двокроковий прогноз, отриманий за МПТ

Результати експерименту наведені в таблиці 3.3. Як видно з таблиці, найкращі результати на ділянці прогнозування показав поліноміальний дворядний МГВА : прогнозоване значення відрізняється від реального лише на 0.7%, що є дуже хорошим результатом, при цьому, варто відмітити, що при відновленні моделі даних метод показав значно гірші результати : СКП для відновленої моделі -37,11; СКП для прогнозу – 8,57.

Таблиця 3.3 – Результати Експерименту 1.

NYSE Composite	модель			прогноз			
	СКП(RMSE)	МАП	КД	СКП(RMSE)	САПП(MAPE)	МіАПП	МАПП(MAE)
ЕЗ	56.740	73.14	0.86	58,84	4,97	3,49	6,45
МГВА	37.110	92.70	0.91	8,57	0,74	0,62	0,87
НМ	52.065	148.83	0.86	68,99	5,86	4,26	7,46
МІТ	–	–	–	25,38	2,22	2,00	2,45

Експеримент 2. В цьому експерименті ми будуватимемо прогноз на 3 кроки для ряду відношень валют євро та долару в період з 1 грудня 2018 року по 19 квітня 2019 року(додаток В). Даний ряд містить 106 спостережень. Для побудови прогнозу методом групового врахування аргументів використано дворядний алгоритм із розміром навчальної вибірки 70% від загального числа спостережень (рисунок 3.8).

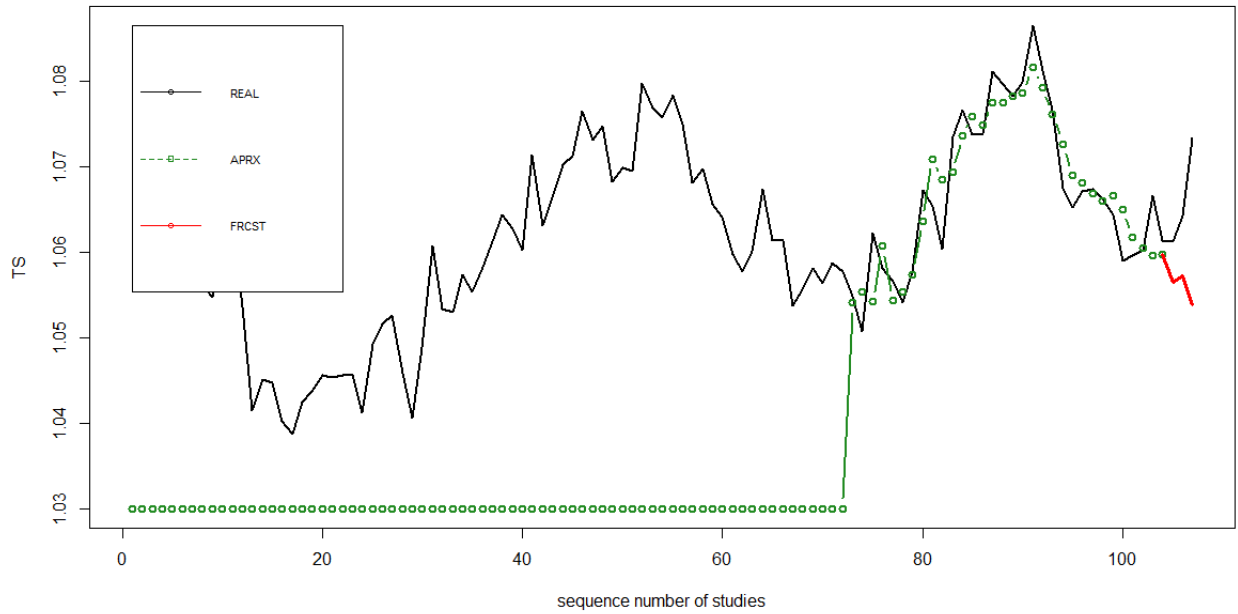


Рисунок 3.8– Прогноз співвідношення EUR\USD МГВА

Графіки прогнозів за допомогою нейронної мережі, методом подібних траєкторій та методом експоненційного згладжування наведені на рисунках 3.9 – 3.11 відповідно.

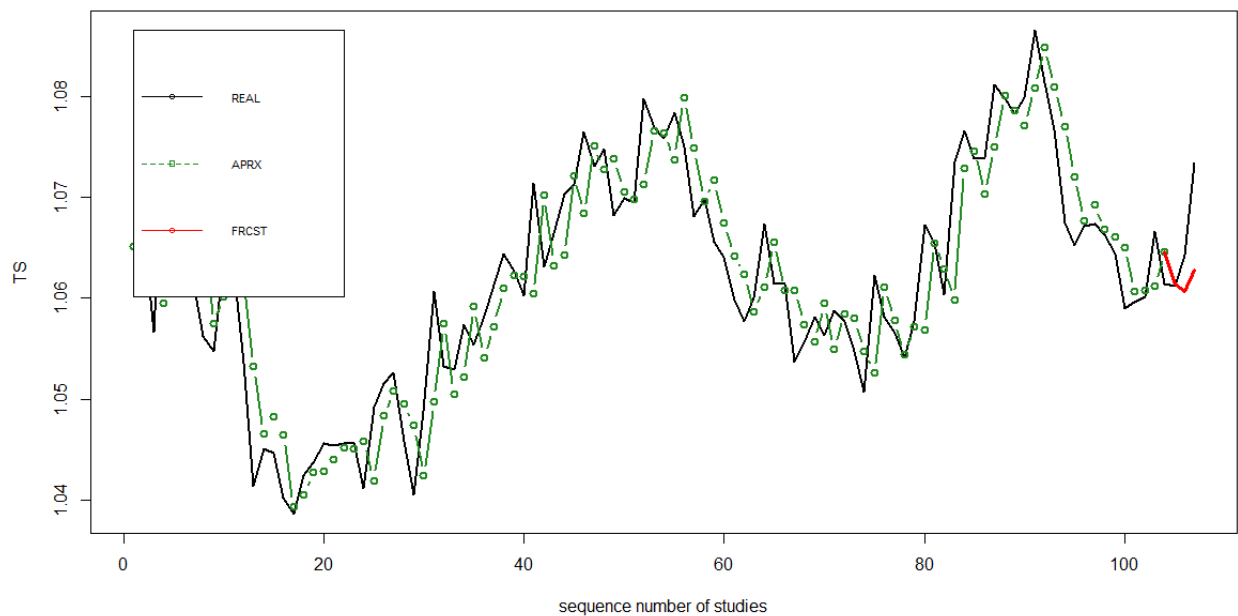


Рисунок 3.9– Прогноз співвідношення EUR\USD за допомогою нейронної мережі

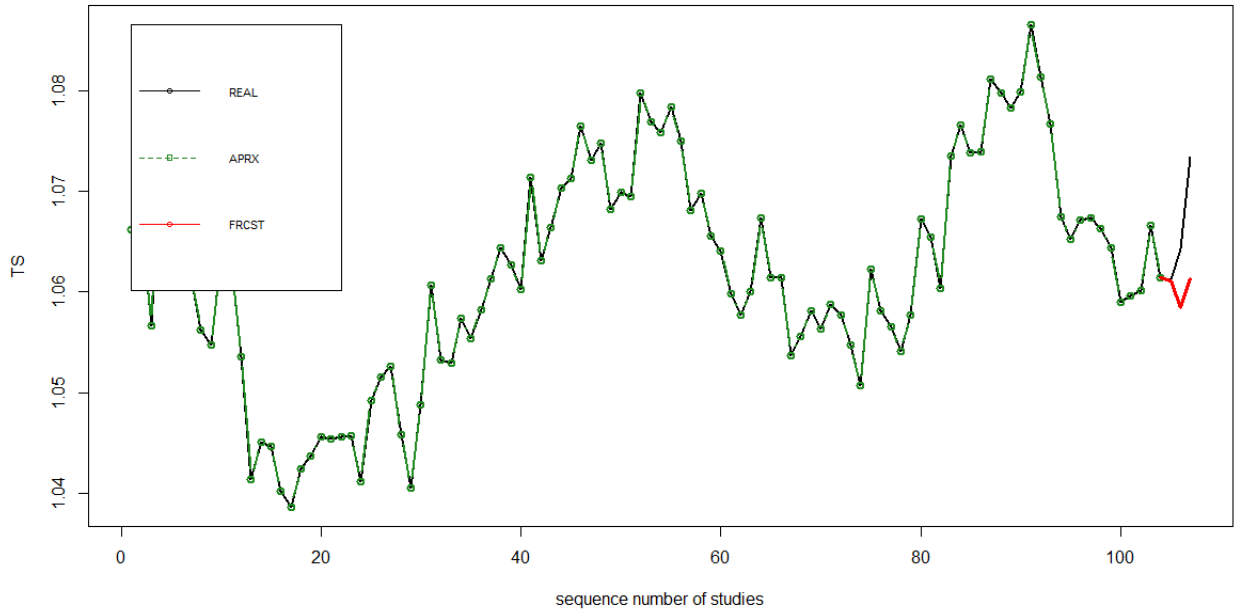


Рисунок 3.10– Прогноз співвідношення EUR\USD методом подібних траєкторій

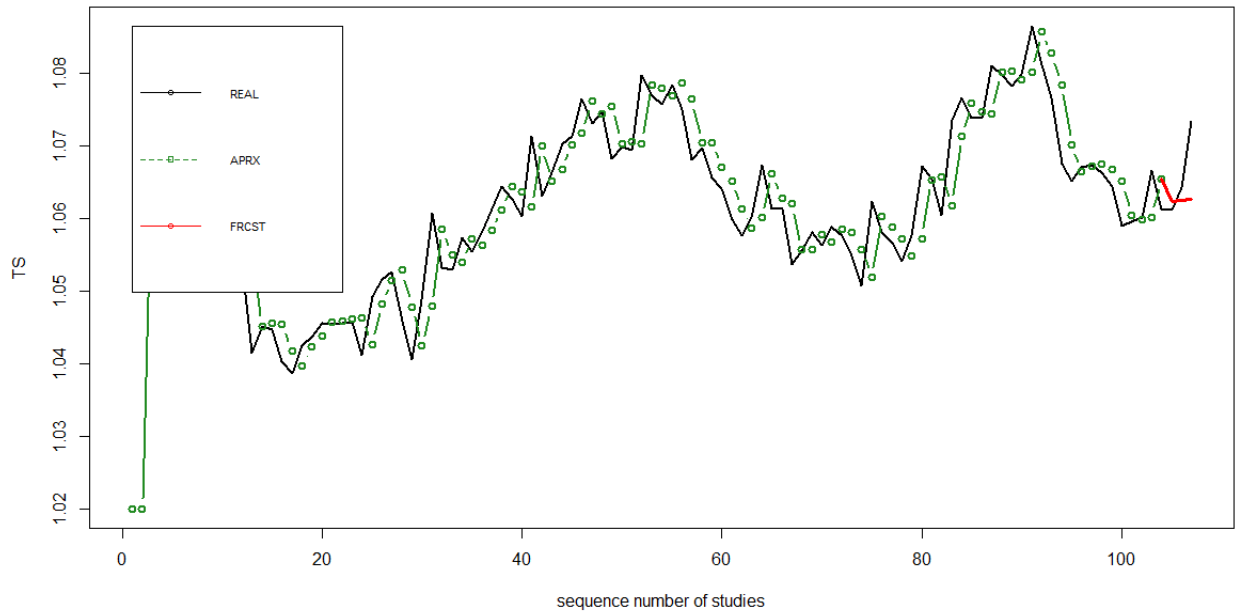


Рисунок 3.11– Прогноз співвідношення EUR\USD методом експоненційного згладжування

Результати проведеного експерименту наведені в таблиці 3.4, з яких видно, що найбільш якісний прогноз отримано методів експоненційного

згладжування та подібних траєкторій. В той же час найгіршим за якістю виявився прогноз, отриманий за допомогою МГУА.

Таблиця 3.4– Результати експерименту 2

EUR_USD	модель			прогноз			
	СКП	МАП	КД	СКП(RMSE)	САПП	МіАПП	МАПП
ЕЗ	0,00611	0,01051	0.84072	0,006341	4,2792265	1,117602	10,03540
МГВА	0,00692	0,01099	0.79006	0,009816	7,8866218	1,850735	13,24669
НМ	0,00599	0,01201	0.97569	0,006480	4,5105235	0,253486	9,904974
МІТ	–	–	–	0,007759	5,6239769	0,091405	11,2623

Експеримент 3.

В цьому експерименті ми будуватимемо прогноз на 3 кроки для ряду цін ф'ючерсів на пшеницю в США . Даний ряд містить 72 спостереження. Для побудови прогнозу методом групового врахування аргументів використано дворядний алгоритм із розміром навчальної вибірки 47 спостережень (рисунок 3.12).

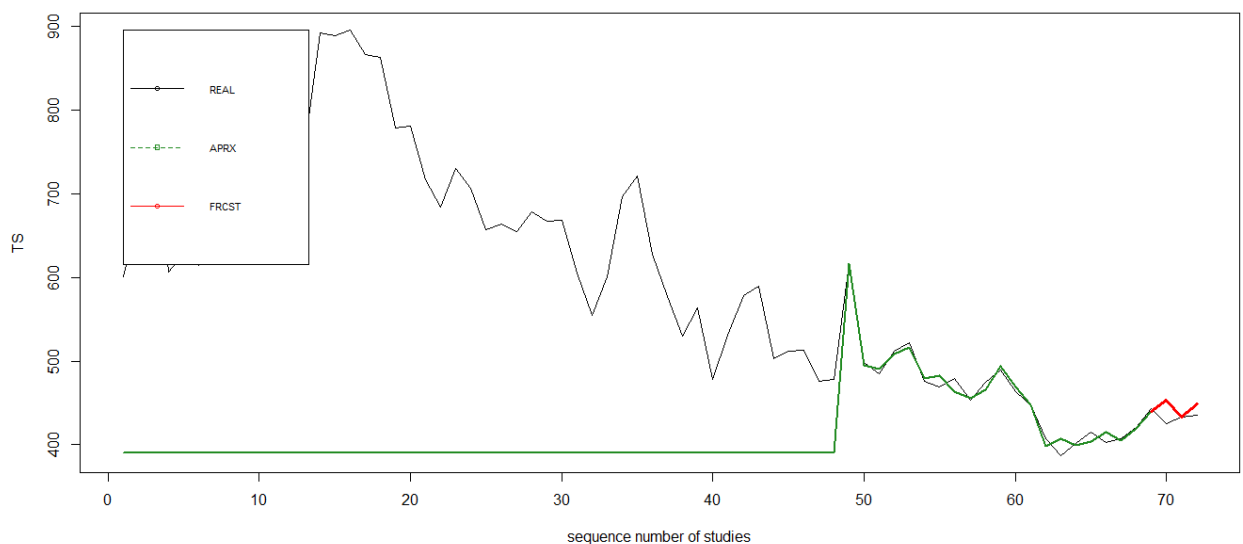


Рисунок 3.12– Прогноз цін ф'ючерсів на пшеницю в США методом ГВА

Графіки прогнозів за допомогою нейронної мережі, методом подібних траєкторій та методом Хольта наведені на рисунках 3.13 – 3.15 відповідно.

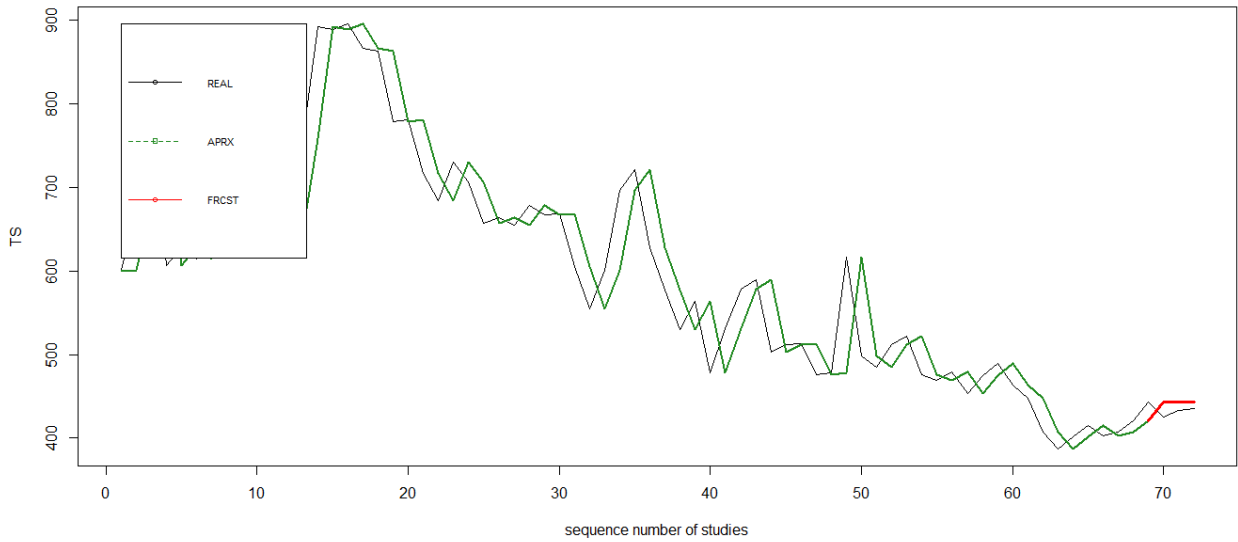


Рисунок 3.13– Прогноз цін ф’ючерсів на пшеницю в США за допомогою нейронної мережі

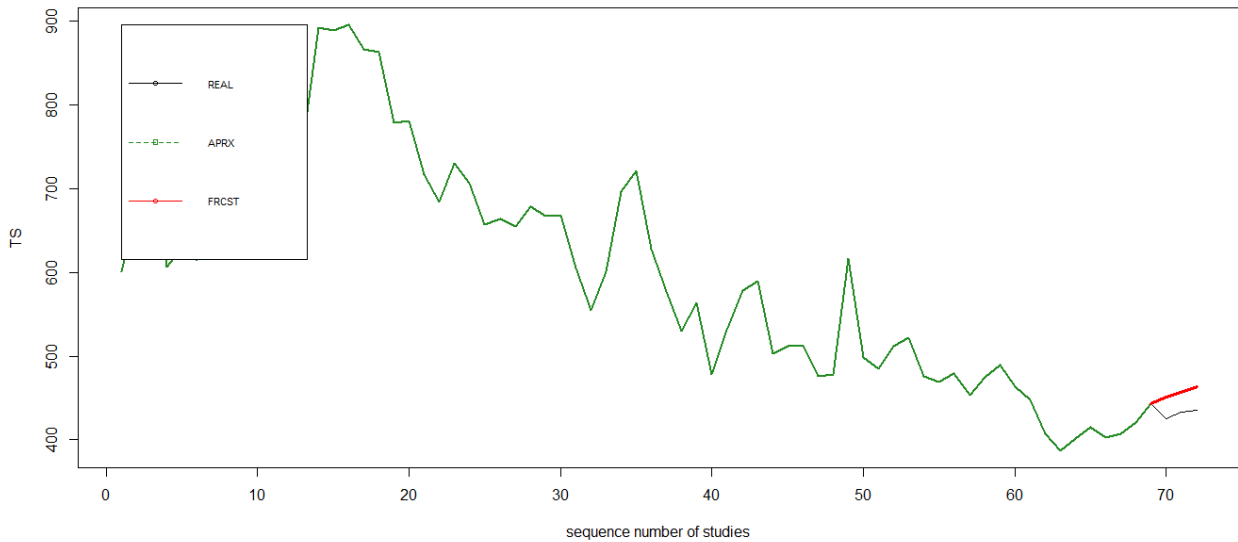


Рисунок 3.14– Прогноз цін ф’ючерсів на пшеницю в США методом подібних траєкторій

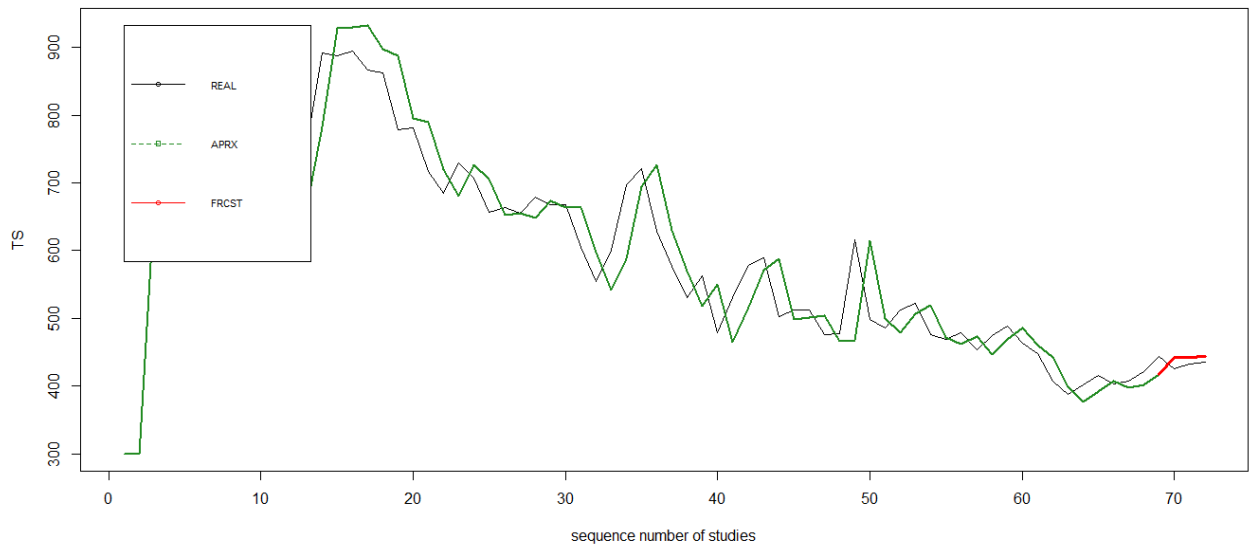


Рисунок 3.15– Прогноз цін ф'ючерсів на пшеницю в США методом ЕЗ

Результати експерименту наведені в таблиці 3.5. Як видно з результатів, найкращий прогноз отримано методом експоненційного згладжування (метод Хольта) та за допомогою НМ. В той час як решта методів виявились менш точними, найгіршим виявився прогноз, отриманий за допомогою методу подібних траєкторій.

Таблиця 3.5– Результати експерименту 3

wheat_price	модель			прогноз			
	СКП	МАП	КД	СКП	САПП	МіАПП	МАПП
ЕЗ	11,6623	19,7324	0,84072	12,201	27,0619	18,257	39,107
МГВА	18,2465	21,407	0,9746	17,950	32,4788	0,978	66,077
НМ	12,0059	12,0120	0,83220	13,114	28,9831	18,943	42,278
МРТ	-	-	-	25,931	59,9979	56,979	64,493

Експеримент 4.

Для побудови мережі Байєса використовувалися п'ять індексів, які описані вище. Вибірка була поділена на навчальну (527 випадків) та перевірочну (50 випадків).

Дослідження зручно проводити, розглядаючи величину відхилення індексу від його попереднього значення [4]. Саме такий підхід використаний в роботі.

$$y_{k,i}^{new} = y_{k,i} - y_{k-1,i},$$

де $i = \underline{1,5}$ – номер ряду, $y_{k,i}$ – історичні дані за період k для ряду i , $k = \underline{2,n}$, де n – кількість досліджень (розмірність вибірки).

Перед початком побудови моделі проводимо процедуру дискретизації даних. В нашому випадку проводиться ієрархічна дискретизація. Кількість проміжків дискретизації обиралась емпіричним шляхом. Результат дискретизації на прикладі одного індексу зображено нижче, на рисунку 3.16.

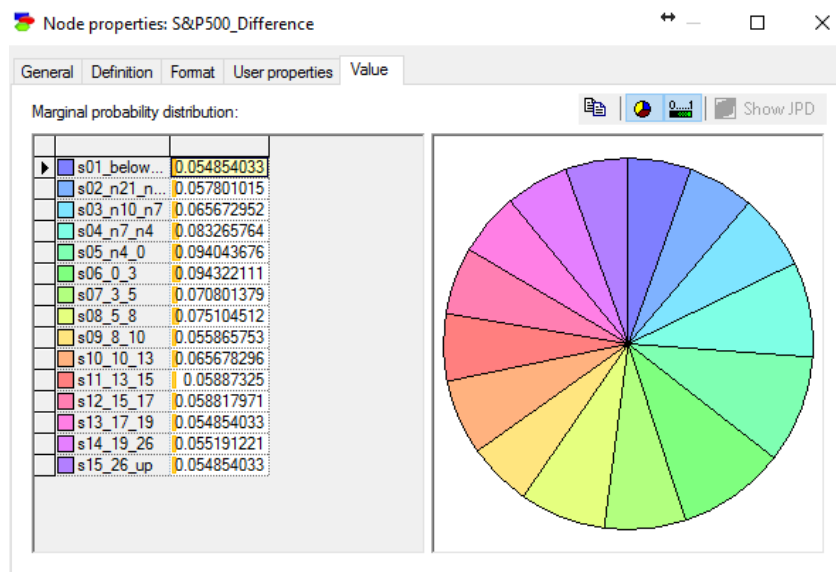


Рисунок 3.16 – Дискретизація значень ряду різниць індексу S&P500

Далі вибирається відповідний алгоритм навчання мережі у разі необхідності використовується експертне знання і виконується навчання мережі.

При побудові структури мережі Байєса в програмі GeNIe слід пам'ятати, що обраний алгоритм впливає на швидкість і на саму побудовану структуру.

Найшвидшим виявився алгоритм Greedy Thick Thinning, його використано для аналізу прикладів. У результаті роботи алгоритму отримуємо тільки одну структуру, яка є логічною і оптимальною (рис 3.17).

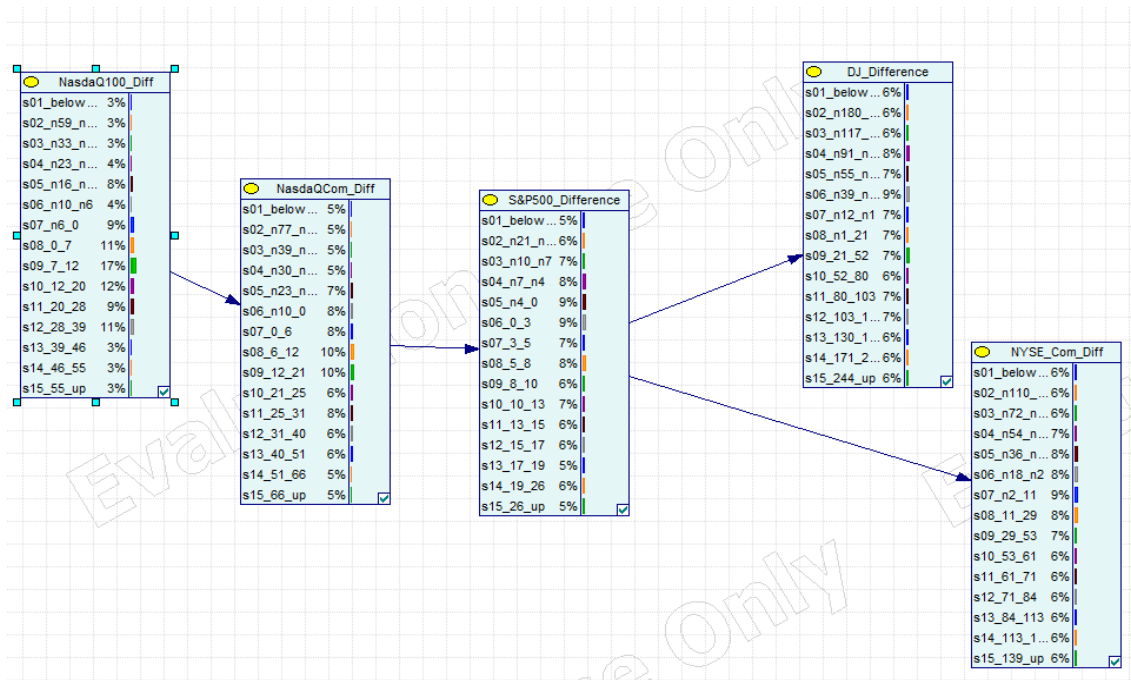


Рисунок 3.17 – Структура мережа Байєса у вигляді гістограм вибірки

По структурі отриманої мережі робимо висновок щодо індексу для прогнозування – Nasdaq100. На рисунку 3.18 бачимо результат одного з експериментів.

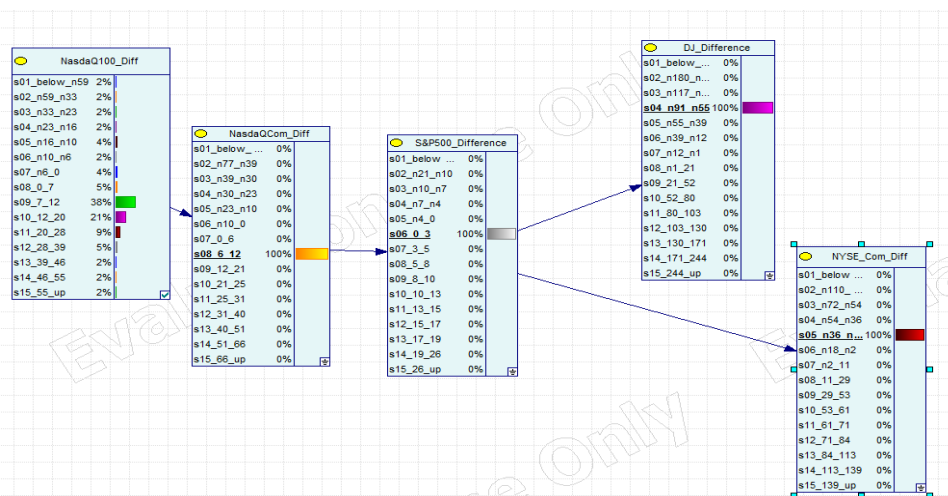


Рисунок 3.18 - Результат прогнозування зміни індексу Nasdaq100

для 19.04.2019р

На рис. 3.19 показана діаграма зміни прогностного і реального відхилень індексу на останніх 50 часових інтервалах.

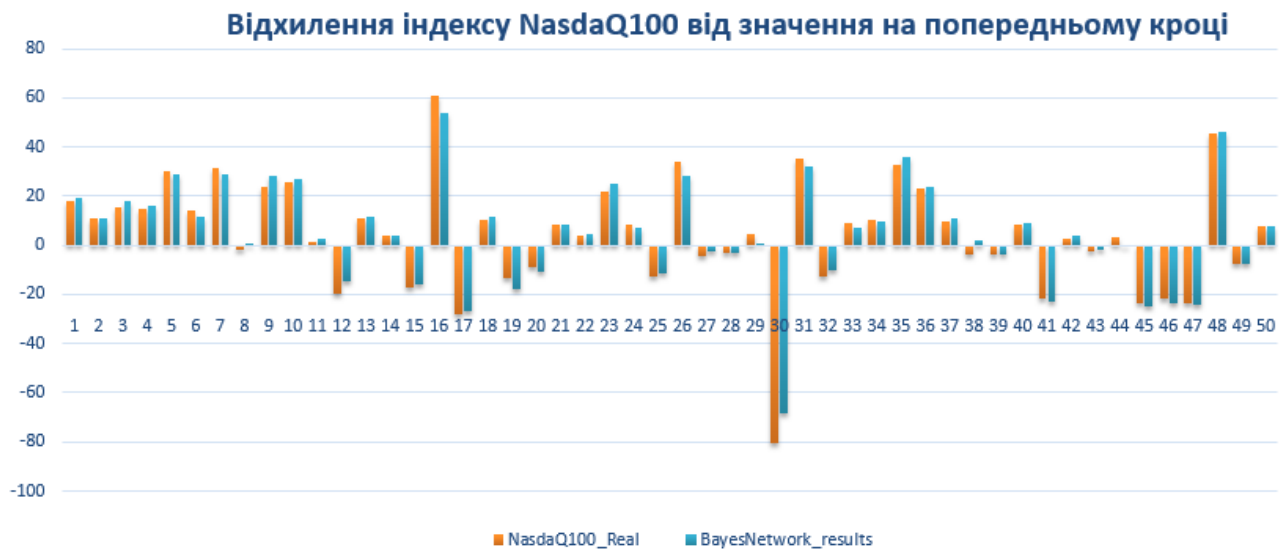


Рисунок 3.19 – Відхилення індексу Nasdaq100 від значення на попередньому кроці

Після аналізу результатів, бачимо, що у 94% випадків було точно передбачене зростання або падіння значення обраного індексу. В 86% випадків прогностне і реальне значення відхилення індексу Nasdaq100 від значення в попередній момент часу мають відхилення менше 10%. При цьому 10% експериментів мають відхилення більше 15% – така ситуація характерна для малих різниць та може бути пов'язана з варіантом дискретизації змінної.

3.4 Висновки до розділу 3

При прогнозуванні фінансово-економічних показників необхідно раціонально обрати метод, який якісно прогнозує необхідний показник. При

цьому слід також визначити параметри метода відповідно до значень ряду, що прогнозується. Таким чином найбільш "вузьким" місцем є вибір методу та параметрів моделі в залежності від особливостей вхідних даних для знаходження найкращого варіанту.

Для порівняння методів, що практично реалізовані в розробленій програмі було використано три часові ряди різної довжини та характеру. Так, перші два ряди мають нестационарний характер та недетермінований тренд. Третій ряд має від'ємний тренд. Тому комплексний аналіз результатів прогнозування цих рядів дає змогу створити широку картину роботи методів, що розглядаються.

Перш за все слід підкреслити, що метод експоненційного згладжування дає прогноз лише на 1 крок, тому його результати інших прогнозів в цьому аналізі враховуватись не будуть. Експоненційне згладжування є найпростішим з розглянутих методів. Цей метод дає достатньо якісний для практичного використання прогноз на 1 крок за умови стаціонарності ряду або чіткої тенденції без суттєвих коливань. Тому область можливого використання метода досить обмежена.

Прогнозування за допомогою нейронної мережі в цілому дало хороший результат. Проте при прогнозуванні на більшу кількість кроків якість суттєво погіршується. Це можна пояснити використанням простої мережі- з одним прихованим шаром.

Метод групового врахування аргументів видає дуже якісні прогнози при короткостроковому прогнозуванні. Також зазначимо, що в порівнянні з усіма іншими методами МГВА значно програє в швидкості роботи, однак дозволяє отримати високоякісну модель при малій кількості вихідних даних, що дозволяє побудувати якісний прогноз.

Метод пошуку подібних траєкторій чудово працює на великих рядах, генеруючи достатньо якісні прогнози навіть при прогнозуванні на велике число кроків. На малих рядах метод погано працює при прогнозуванні рядів,

що мають значні коливання (тобто при неможливості знайти достатньо близьку траєкторію).

При побудові мережі Байєса отримали, що 86% випадків прогнозу і реального значення відхилення індексу Nasdaq100 від значення в попередній момент часу мають відхилення менше 10%. При цьому, при прогнозуванні на більшу кількість кроків якість прогнозу не погіршується. Для покращення результату варто звернути увагу на вибір кількості проміжків дискретизації та, за можливості, збільшити навчальну вибірку для мережі.

Таким чином, використані методи прогнозування дають можливість отримати якісний короткостроковий прогноз, про що свідчать отримані результати моделювання. Всі статистичні методи прогнозування зведені в єдиний модуль, що є зручним інструментом.

РОЗДІЛ 4 РОЗРОБКА СТАТАП-ПРОЕКТУ

4.1 Інформаційна карта проекту

В таблиці 4.1 показана інформаційна карта стартап-проекту. Це стандартний спосіб ознайомлення з ідеєю проекту для замовників, спонсорів та інших зацікавлених сторін. Є дуже важливою для заповнення, так як складає перше враження про розроблюваний продукт.

Таблиця 4.1 – Інформаційна карта проекту

Назва проекту	SmartForecast
Анотація	Сьогодні практично в усіх сферах використовують ПК. І щорічно компанії витрачають шалені гроші на ліцензійні програмні продукти. Проте зібрана командою статистика показала, що майже третина з них не використовується. Продукт SmartForecast прогнозує оптимальну кількість різних програмних продуктів, яку потрібно купувати компанії для забезпечення працівників лише справді потрібними програмами. Також розраховуються прогнози щодо росту\спаду необхідної кількості програмних продуктів у майбутньому при різних поведінках ринку, росту\зниження цін на акції компанії, виходу її на нові ринки, тощо.
Термін реалізації	12 місяців
Необхідні ресурси	Фінансові : <ul style="list-style-type: none"> – 200\$ -електроенергія; – маркетингові дослідження.

Продовження таблиці 4.1

Необхідні ресурси	<p>Матеріальні:</p> <ul style="list-style-type: none"> – персональні комп’ютери та їх комплектуючі – 2шт; – МФУ – 1шт; – папір для друку – 2пачки. <p>Людські:</p> <ul style="list-style-type: none"> – програміст – 1; – дизайнер – 1; – менеджер – 1.
Опис проблеми, яку вирішує проект	<p>Компанії витрачають значні кошти для купівлі ліцензійного програмного забезпечення, при цьому, забезпечуючи працівників продуктами, які їм не потрібні. Така ситуація частіше всього виникає у великих компаніях та корпораціях, де, зазвичай, відбувається постійна зміна кадрів. Проект вирішує проблему нераціональних витрат коштів компанії на програмні продукти для ПК працівників.</p>
Головні цілі та завдання проекту	<p>Метою проекту є розрахунок оптимальної кількості замовлень ліцензійних програмних продуктів різного призначення для економії коштів компанії</p>
Очікувані результати	<p>Компанії витратимуть гроші лише на ту кількість програмних продуктів, яка необхідна на даному етапі. За результатами проведеної статистики, це дозволить зекономити близько 10% щорічних витрат компанії. В довгостроковій перспективі, планується розширити можливості створюваного продукту для прогнозування потрібних потужностей персональних комп’ютерів, а також росту\зниження цін на акції компанії, виходу її на нові ринки, тощо.</p>

4.2 Команда стартап-проекту

В таблиці 4.2 перелічені функціональні обов'язки виконавців проекту.

Таблиця 4.2 - Опис функціональних обов'язків виконавців проекту

Посада	Функціональні обов'язки
Керівник проекту, менеджер	<ul style="list-style-type: none"> – планування бюджету, трудовитрат – побудова ієрархічної структури робіт проекту; – оформлення результатів роботи; – ведення проектної документації; – аналіз тендерних пропозицій; – управління командою, координація взаємодії з замовниками і спонсорами проекту; – розподіл задач між членами команди; – аналіз ризиків.
Програміст	<ul style="list-style-type: none"> – розробка логічної моделі та архітектури проекту; – написання вихідних кодів програм, інтеграція зовнішніх даних та продуктів; – написання тест-плану та тест-кейсів, тестування продукту на всіх етапах розробки; – захист інформації.
Дизайнер	<ul style="list-style-type: none"> – визначити цільову аудиторію майбутнього сайту; – спланувати структуру сайту; – розробити схему навігації та інтерфейсу; – налагодити якісну взаємодію користувача з системою.

4.3 Маркетингова стратегія та маркетинговий план проекту

Короткий опис ідеї , його основні напрямки застосування, а також вигоди для користувачів проекту описані в таблиці 4.3.

Таблиця 4.3 - Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
Розробка програми, яка, за наданою історичною інформацією виконує	Великі компанії та корпорації	Економія річного бюджету, збільшення ефективності працівників за рахунок обмеження доступу до непотрібних для роботи продуктів тощо.
прогнозування фінансово-економічних процесів, вирішує проблему вибору кращого методу прогнозування та будує оцінки прогнозу.	Середній та малий бізнес	Оптимізація поточних витрат на програмні продукти, планування майбутніх витрат за рахунок обчислених прогнозів

Запропонований продукт являється потенційно новим на ринку та не має замінників. Конкурентами можуть виступати потенційні замовники – великі компанії та корпорації, які, маючи достатньо коштів, можуть дозволити собі розробити товар-замінник. Далі визначимо сильні та слабкі сторони проекту в таблиці 4.4.

Таблиця 4.4 - Визначення сильних, слабких та нейтральних характеристик ідеї проекту

Техніко-економічні характеристики	Потенційні товари та концепції конкурентів		Слабка сторона	Нейтральна сторона	Сильна сторона
	Мій проект	Конкурент			
Ціна	1500\$	3000\$			Доступність на ринку
Підтримка	2 роки	2 роки		За додаткову плату час підтримки продукту може змінюватись	
Продовження таблиці 4.4 Програмні ресурси	Open-source ресурси	Платні ресурси	Час на написання частин, які є реалізовані в платних продуктах		Фінансова і технічна незалежність від обраного ресурсу
Грошові ресурси	800\$	>>800\$	Більші можливості конкурентів		

Далі проведемо аудит технології, за допомогою якої можна реалізувати ідею проекту (технології створення товару). Визначення технологічної здійсненності ідеї проекту передбачає аналіз таких складових (таблиця 4.5) [17]:

- за якою технологією буде виготовлено товар згідно ідеї проекту?

- чи існують такі технології, чи їх потрібно розробити/доробити?
- чи доступні такі технології авторам проекту?

Таблиця 4.5 - Технологічна здійсненність ідеї проекту

Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій
працівниками. Розраховуються прогнози щодо росту\спаду необхідної кількості програмних продуктів у майбутньому	Програмне забезпечення: <ul style="list-style-type: none"> – Visual Studio 2015; – C#; – MS SQL Server 2016; 		
	Апаратне забезпечення: <ul style="list-style-type: none"> – HDD 300 – CPU INTEL CORE I5+ – ОЗУ 8gb 		
	Програмне забезпечення: <ul style="list-style-type: none"> – Visual Studio 2015; – Python; – MySQL. Апаратне забезпечення: <ul style="list-style-type: none"> – ОЗУ 8gb; – HDD 300; – CPU INTEL CORE I5+. 	Наявні	Доступні

Обрана технологія реалізації ідеї проекту: ООП, Visual Studio 2017, C#, MS SQL Server 2018 – доступна та наявна технологія на ринку, обізнаність команди в технологіях, можливість складних обчислень.

На наступному кроці проаналізуємо ринкові можливості запуску стартап-проекту. Визначення ринкових можливостей, які можна використати під час ринкового впровадження проекту, та ринкових загроз, які можуть перешкодити реалізації проекту, дозволяє спланувати напрями розвитку проекту із урахуванням стану ринкового середовища, потреб потенційних клієнтів та пропозицій проектів-конкурентів [14, 18]. Спочатку проводиться аналіз попиту (таблиця 4.6):

- наявність попиту;
- обсяг;
- динаміка розвитку ринку.

Таблиця 4.6 - Попередня характеристика потенційного ринку стартап-проекту

Показники стану ринку	Характеристика
Кількість головних гравців, од	~1000
Загальний обсяг продаж	€1,7 млрд.
Динаміка ринку (якісна оцінка)	ринок зростає
Наявність обмежень для входу (характер обмежень)	відповідність продукту стандартам якості
Специфічні вимоги до стандартизації та сертифікації	щорічна сертифікація якості продукту та його модифікацій
Середня норма рентабельності в галузі (або по ринку)	15%

Як бачимо з таблиці, ІТ ринок в Україні стрімко розвивається і є привабливим для входження за попереднім оцінюванням.

Надалі визначаються потенційні групи клієнтів, їх характеристики, та формується орієнтовний перелік вимог до товару для кожної групи (таблиця 4.7).

Таблиця 4.7 - Характеристика потенційних клієнтів стартап-проекту

Потреба, що формує ринок	Цільова аудиторія	Відмінності у поведінці різних потенційних груп клієнтів	Вимоги споживачів до товару
економія витрат компаній на ПЗ працівників	великий і середній бізнес. Рідше – малий бізнес	фінансові можливості та причини купівлі товару	наявність сертифікатів якості, патентів

Після визначення потенційних груп клієнтів проводимо аналіз ринкового середовища: складаємо таблиці факторів, що сприяють ринковому впровадженню проекту, та факторів, що йому перешкоджають (таблиці 4.8 – 4.9). Фактори в таблиці подані в порядку зменшення значущості.

Таблиця 4.8 - Фактори загроз

Фактор	Зміст загрози	Можлива реакція компанії
Нестабільність у світі	Компанії боятимуться витратити кошти на продукти не першої необхідності	Аналіз ринку та пошук нових клієнтів, які залишаються відносно стабільними, модифікація продукту під потреби ринку, рекламна кампанія висвітлює ідеї економії річного бюджету.
Незадовільна економічна ситуація в державі	Можливе зменшення розмірів та бюджету компаній, що загрожує зменшенням попиту на продукт	рекламна кампанія висвітлює ідеї економії річного бюджету

Продовження таблиці 4.8

Поява конкурентів	Зменшення долі ринку Поява товарів-аналогів чи замінників за нижчою ціною, відомішим брендом, новими можливостями тощо, підвищення бар'єру входження на ринок	Створення бонусної (наприклад) системи для існуючих клієнтів, підвищення якості продуктів
Нові вимоги щодо сертифікації продуктів	Нові сертифікації значно підвищать ціну на продукт або взагалі будуть недоступні для проходження	Пошук нових спонсорів

Таблиця 4.9 - Фактори можливостей

Фактор	Зміст можливості	Можлива реакція компанії
Зростання попиту на ІТ товари	Ріст попиту призведе до росту існуючих компаній та створення нових, що в свою чергу збільшить попит на використання продукту. Залучення нових спонсорів та інвесторів	Робота по залученню нових спонсорів та інвесторів, розширення компанії, розширення рекламної кампанії, за потреби, модифікація продукту під потреби ринку
Нові ринки збуту	Вихід продукту на міжнародний рівень	Переклад продукту на мови потенційних країн-замовників
Зміна курсу \$	Збільшення попиту на продукт в аутсорсингових компаніях, як наслідок - нові клієнти. Або зріст попиту на продукт в українських продуктових компаніях	Вивчення ніші компаній-замовників, розширення рекламної кампанії, перегляд умов продажу продукту продуктовим українським компаніям (бонуси, зменшення вартості тощо)
Створення нових ІТ-технологій	Можливість росту ІТ-ринку та захоплення ним нових ніш	Аналіз можливих нових сфер діяльності для ІТ-компаній. Модифікація продукту під потреби ринку, рекламна кампанія для кожного сегменту ринку

Продовження таблиці 4.9

Незадовільна економічна ситуація в державі	Виїзд кваліфікованих кадрів за кордон, що призводить до зменшення розмірів та бюджету компаній. Це спонукає компанії до більш грамотного розподілу бюджету	Рекламна кампанія направлена на висвітленні ідеї, що за рахунок нашого продукту компанія зможе зекономити до 12% річного бюджету, що в нестабільній економічній ситуації є вагомим фактором
Незадовільна економічна ситуація в державі	Виїзд кваліфікованих кадрів за кордон, що призводить до зменшення розмірів та бюджету компаній. Це спонукає компанії до більш грамотного розподілу бюджету	Рекламна кампанія направлена на висвітленні ідеї, що за рахунок нашого продукту компанія зможе зекономити до 12% річного бюджету, що в нестабільній економічній ситуації є вагомим фактором
Скорочення попиту на ІТ-товари	Зменшення розмірів та бюджету компаній. Це спонукає компанії до більш грамотного розподілу бюджету	Рекламна кампанія направлена на висвітленні ідеї економії річного бюджету, що в нестабільній економічній ситуації є вагомим фактором

Надалі проводиться аналіз пропозиції: визначаються загальні риси конкуренції на ринку (таблиця 4.10) [17].

Таблиця 4.10 - Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства
Тип конкуренції - монополія	Можлива поява компаній, які перешкоджатимуть продажу продукту за рахунок свого бренду, налагодженій структурі клієнтів тощо	Залучення клієнтів за рахунок ціни продукту, аналіз потреб ринку та розширення функцій продукту для їх задоволення, пошук нових спонсорів, аналіз конкретного клієнта та

Продовження таблиці 4.10

		підлаштування умов співпраці з ним
За рівнем конкурентної боротьби - міжнародний	Ринок ІТ сьогодні майже не знає бар'єрів, тому поява аналогічного продукту в іншій країні не залишиться непоміченою	Постійний моніторинг нових товарів, компаній, аналіз ринку та економічної ситуації у світі та прогнозування можливостей появи нових конкурентів
За галузевою ознакою - міжгалузева	Конкурентами можуть виступати лише ІТ-компанії чи ІТ-стартапи	Моніторинг динаміки ІТ-ринку, аналіз ризиків
Конкуренція за видами товарів - товарно-видова	Можлива проява товарів-замінників чи аналогічних продуктів з додатковими можливостями, зручнішим інтерфейсом тощо	Розширення команди аналітиком (маркетологом) – аналіз потреб клієнтів, їх побажань тощо. Розвиток команди для можливості використання більш нових, складних технологій
За характером конкурентних переваг - нецінова	Використання конкурентами нових дорогих технологій, інновацій, бренду тощо	Залучення клієнтів за рахунок ціни продукту, аналіз потреб ринку та розширення функцій продукту, пошук нових спонсорів, аналіз конкретного клієнта та підлаштування умов співпраці з ним
За інтенсивністю - немарочна	Поява на ринку аналогічних чи схожих продуктів відомих брендів. Поява конкурента, з непопулярним брендом, проте з ширшим спектром функцій, використанням новіших технологій тощо	Залучення клієнтів за рахунок ціни, аналіз потреб ринку та розширення функцій продукту для їх задоволення, пошук нових спонсорів, налаштування умов співпраці з клієнтами

Для більш детально планування, після аналізу конкуренції варто провести більш детальний аналіз умов конкуренції в галузі за моделлю М. Портера.

4.4 Елементи фінансової моделі

В таблиці 4.11 наведено основні сукупні інвестиційні витрати, потрібні на початкових етапах реалізації стартап-проекту.

Таблиця 4.11 - Інвестиційні витрати на реалізацію стартап-проекту

Стаття витрат	Сукупні витрати
Проведення пошукових та прикладних досліджень	0
Розробка проектних матеріалів і ТЕО	600\$
Робоче проектування і прив'язка проекту	0
Витрати на придбання обладнання та устаткування та пристроїв	430
Витрати на придбання нематеріальних активів	0
Витрати на утримання обладнання та приміщень	0
Витрати на передвиробничі маркетингові дослідження	0
Витрати на створення збутової мережі	0
Витрати на просування та рекламу	70\$
Оплата юридичних послуг	110\$
Витрати на матеріальні ресурси	300\$
Витрати на оплату праці команди стартапу	1000\$
Разом:	2510\$

В таблиці 4.12 визначені основні фінансово-економічні показники проекту.

Таблиця 4.12 - Визначення основних фінансово-економічних показників проекту

Стаття витрат	Сукупні витрати
Обсяг виробництва продукції в натуральних показниках	5
Собівартість одиниці продукції	3410\$
Собівартість виробництва продукції	17050\$
Обсяг реалізації продукції в натуральних показниках	5
Ціна реалізації продукції без ПДВ	4150\$
Виручка від реалізації продукції без ПДВ	20750\$
Податок на додану вартість (ПДВ)	4150\$
Валовий прибуток	3700\$
Податок на прибуток	740\$
Чистий прибуток	2960\$

Рентабельність продажів (або маржа прибутку) показує, скільки прибутку приносить кожна гривня з обсягу реалізації. Маржу прибутку, як правило, визначають окремо за кожним видом діяльності або за кожною групою реалізованої продукції за формулою:

$$R_s = \frac{П}{В} 100\% = 0.142 * 100 = 14,2\% \quad (4.1)$$

де П – прибуток; В– виручка від реалізації продукції.

Період окупності проекту відображає час, який потрібен для того, щоб сума надходжень від реалізації проекту відшкодувала суму витрат на його впровадження. Період окупності звичайно вимірюється в роках або місяцях та може бути розрахований за формулою :

$$PBP = \frac{\Pi}{ACI} = 2410 \backslash 2960 = 0,82 \text{ року} \quad (4.2)$$

де PBP – період окупності інвестицій, роки; Π – сума інвестиційних витрат; ACI – щорічні надходження (річний чистий прибуток).

Рентабельність довгострокових інвестицій – коефіцієнт повернення інвестицій, показник рентабельності вкладень, що у відсотковому співвідношенні демонструє прибутковість (у разі, коли його значення більше 100%) або збитковість (у разі, коли його значення менше 100%) інвестицій в проект. Якщо розрахований показник менший 100%, то інвестиційні вкладення не окупаються. Показник може бути розрахований за наступною формулою:

$$ROI = \frac{D-C}{\Pi} 100\% = \frac{2960}{2510} 100\% = 118\% \quad (4.3)$$

де D – дохід (виручка від реалізації продукції); C – повна собівартість; Π – сума інвестиційних витрат [14].

Бачимо, що рентабельність проекту становить 118%, що є дуже хорошим показником.

За результатами аналізу потенційних груп споживачів (сегментів) автори ідеї обирають цільові групи, для яких вони пропонуватимуть свій товар, та визначають стратегію охоплення ринку - наша компанія працює із всім ринком, пропонуючи стандартизовану програму (включно із характеристиками товару/послуги) – вона використовує масовий маркетинг.

4.5 Стратегія конкурентної поведінки

Наступним кроком є вибір стратегії конкурентної поведінки. Для даного проекту було обрано стратегію лідера, обґрунтуємо це.

Залежно від міри сформованості товарного(галузевого) ринку, характеру конкурентної боротьби компанії-лідери обирають одну з трьох стратегій: розширення первинного попиту, оборонну або наступальну стратегію або ж застосувати демаркетинг або диверсифікацію [13].

Стратегія розширення первинного попиту доцільна у разі, якщо фірмі-лідерові недоцільно розмінюватися на боротьбу з невеликими конкурентами, вона може отримати велику економічну віддачу від розширення первинного рівня попиту. В цьому випадку компанія займається реалізацією заходів по формуванню попиту(навчанню споживачів користуванню товаром, формування регулярного попиту, збільшення разового споживання), також пропаганду нових напрямів застосувань існуючих товарів, виявлень нових груп споживачів. Розширюючи таким чином ринковий попит, лідер надає допомогу усім підприємствам, що «йдуть за ним», несучи при цьому основні фінансові витрати, проводячи найбільш революційні НДДКР. Така стратегія можлива тільки на початкових стадіях життєвого циклу товару, коли попит ще є розширюваним, а взаємний тиск конкурентів ще невеликий. Інакше фірмі лідерові необхідно приймати оборонну або наступальну стратегію.

У міру зростання ринку, його становлення позиції компанії-новатора починають атакувати конкуренти-імітатори. В цьому випадку, компанія може вибрати оборонну стратегію, метою якої є захист власної ринкової долі. Оборона може бути:

- інновації з метою постановки технологічних бар'єрів для входу в ринок нових конкурентів, подальшого збільшення відриву від них;

- ліквідація ніш для проникнення конкурентів за допомогою розширення товарного асортименту, цінових парасольок, захоплення каналів збуту;
- ведення цінової війни і/або проведення масованої рекламної атаки.

Наступальна стратегія припускає збільшення своєї частки ринку. При цьому переслідувана мета полягає в подальшому підвищенні прибутковості роботи компанії на ринку за рахунок максимального використання ефекту масштабу. Проте, існує межа, при перевищенні якої подальше зростання частки ринку стає невигідним. Це або чисто економічна недоцільність відвойовування добре захищених часток, що сильно захищаються, у дрібніших виробників або ж попадання під дію антимонопольного законодавства [18].

Наступальна стратегія припускає активну інноваційну політику компанії. Вона постійно атакує власні ж досягнення, збільшуючи розрив між собою і основними конкурентами. Постійні техніко-економічні вдосконалення, модифікація розміру і форми упаковки, використання event-маркетингу – типові складові арсеналу фірм-лідерів [16].

Якщо фірма потрапляє під дію антимонопольного законодавства, вона може удатися до стратегії демаркетинга, що припускає скорочення своєї частки ринку, зниження рівня попиту на деяких сегментах за рахунок підвищення ціни. При цьому ставиться завдання недопущення на ці сегменти конкурентів, а компенсація втрат прибутку через зменшення обсягів виробництва компенсується встановленням надвисоких цін.

Проте у більшості випадків найпривабливішою стратегією для компаній-лідерів є диверсифікація, що дозволяє використати переваги масштабу виробництва [15].

4.6 Висновки до розділу 4

Сьогодні важко знайти сферу діяльності, де б не використовувались комп'ютери. І щорічно компанії, особливо великі корпорації, витрачають шалені гроші на ліцензійні програмні продукти та надмірні потужності ПК. Проте зібрана статистика показала, що майже третина з них не використовується. Продукт SmartForecast прогнозує оптимальну кількість різних програмних продуктів, яку потрібно купувати компанії для забезпечення працівників лише справді потрібними програмами. Також розраховуються прогнози щодо росту\спаду необхідної кількості програмних продуктів та ПК у майбутньому при різних поведінках ринку, росту\зниження цін на акції компанії, виходу її на нові ринки, тощо.

Сьогодні на ринку не існує товарів-замінників чи товарів зі схожим функціоналом. Тому обрана стратегія поведінки на ринку – стратегія лідера. Розробка описаного проекту є чудовим способом зайняти нішу на ринку.

Для розробки прототипів достатньо мати команду з трьох людей – менеджера, розробника та дизайнера. Загальні інвестиційні витрати на реалізацію – 2410\$, що окупляться за 10 місяців. Рентабельність інвестицій становить 118%, що є хорошим показником. Рентабельність продажів становить 14,2%. Показник є дуже близьким до рентабельності ІТ-сфери в Україні на сьогодні (15%) - подальша імплементація проекту є доцільною.

Розробка бізнес-плану виконувалась по стандартному шаблону бізнес-планів для стартап-проектів.

ВИСНОВКИ ПО РОБОТІ ТА ПЕРСПЕКТИВИ ДОСЛІДЖЕНЬ

Отже, як вже було сказано, для прийняття коректних, грамотних рішень зокрема в економіці та фінансах необхідно використовувати сучасні методи математичного моделювання фінансово-економічних процесів та чітко формувати альтернативи на основі оцінок прогнозів, отриманих за побудованими моделями. Досягти цього можна за допомогою якісного прогнозування, яке б враховувало всі особливості процесу, а також впливи випадкових збурень. Реалізувати це зручно за допомогою створення інтелектуальних комп'ютерних систем, які значно полегшують якісну обробку інформації для ОПР. Такі програми також дозволяють зекономити багато часу, що покращує ефективність праці ОПР.

В роботі виконано огляд найтипівіших фінансово-економічних процесів та популярних сьогодні статистичних та ймовірнісних методів прогнозування часових рядів. Показано актуальність та важливість якісного розв'язку загальної задачі прогнозування сьогодні. Проведений детальний аналіз наступних методів прогнозування: експоненційне згладжування, метод групового врахування аргументів, нейронні мережі, метод подібних траєкторій та досліджено мережу Байєса. Виконано порівняння оцінок прогнозів.

Для удосконалення роботи та можливостей подальших досліджень необхідно реалізувати й інші методи прогнозування та моделювання часових рядів, виконати більше експериментів для різнотипних рядів, а також допрацювати СППР - розробити зручний інтерфейс користувача та покращити якість обробки вхідних даних для подальшої можливості комерційного застосування системи.

Перспективою для подальших досліджень є покращення точності роботи існуючих методів, їх модифікація, комбінування оцінок прогнозів,

розширення класу процесів, до яких їх можна застосовувати та збільшення горизонту прогнозування.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

- 1 Бідюк П.І. Аналіз часових рядів / П.І. Бідюк, В.Д. Романенко, О.Л. Тимошук (навчальний посібник) — К.: Політехніка, 2010. — 317 с.
- 2 Побудова і застосування байєсівських мереж [Електронний ресурс] / П. І. Бідюк, О. М. Терентьев, О. Л. Жиров, О. В. Гавриленко // Економіка: теорія та практика. - 2014. - № 1. - С. 89-99. - Режим доступу: http://nbuv.gov.ua/UJRN/econom_2014_1_17
- 3 Niedermayer D. An Introduction to Bayesian networks and their contemporary applications [Електронний ресурс] / Innovations in Bayesian Networks — 2008. — 13p. — Режим доступу: <https://www.semanticscholar.org/paper/An-Introduction-to-Bayesian-Networks-and-Their-Niedermayer/264400bc9c967de28fd65aabf744a946aеba8a7a>
- 4 Бидюк П.И. Прогнозирование экономических показателей при помощи сетей Байеса / П.И. Бидюк, А.Н. Терентьев, Л.А. Коршевнюк / Автоматика-2007: сб. науч. работ за материалами XIV междунар. конф., 10-14 вер. 2007 г., Севастополь. — Севастополь: СТУАСТ, 2007. — Т. 2 — С. 91–93.
- 5 Бокс Дж. Анализ временных рядов, прогноз и управление / Дж. Бокс, Г. Дженкинс — М.: Мир, 1974. — 406 с.
- 6 Демиденко Е.З. Линейная и нелинейная регрессия / Демиденко Е.З. — М.: Наука, 1987. — 248с.
- 7 Дрейпер Н. Р. Прикладной регрессионный анализ / Н. Р. Дрейпер, Г. Смит. — К.: Диалектика, 2007. — 912 с.
- 8 Згуровский М. З. Аналитические методы калмановской фильтрации / М. З. Згуровский, В. Н. Подладчиков— К: Наук. думка, 1995. — 285 с
- 9 Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем / Ивахненко А.Г. — К: Наукова думка, 1982. — 296 с.

- 10 Ивахненко А.Г. Обзор задач, решаемых по алгоритмам Метода Группового Учета Аргументов (МГУА) [Электронный ресурс].Режим доступа: <http://www.gmdh.net/articles/rus/obzorad.pdf>
- 11 Ивахненко А.Г. Помехоустойчивость моделирования / А. Г. Ивахненко, В. С. Степашко - К.: Наук. думка, 1985. – 216 с.
- 12 Кулявець В. О. Прогнозування соціально-економічних процесів/ В.О. Кулявець . – К: Кондор, 2009. – 194 с
- 13 Левченко Л.О. Огляд методів прогнозування фінансового стану підприємства на основі економетричних моделей / Л.О. Левченко, Д.С. Белова // Управління розвитком складних систем. - 2013. - Вип. 14. - С. 164-169.
- 14 Льюис К.Д. Методы прогнозирования экономических показателей / Льюис К.Д. ; пер. с англ. и предисл. Е.З. Демиденко. – М.: Финансы и статистика, 1986. – 133 с.
- 15 Хайкин С. Нейронные сети: полный курс. / Хайкин С.– М.: ООО «И. Д. Вильямс», 2006. –1104 с
- 16 Ханк Д.Э. Бизнес-прогнозирование, 7-е издание. / Ханк Д.Э., Уичерн Д., Райс А. Дж. ; пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 656 с.
- 17 Економетрика : Лекції / В. В. Ярова, А. І. Александрова; Харк. нац. аграр. ун-т ім. В.В.Докучаєва. - Х., 2003. - 123 с.
- 18 Chambers J. Software for Data Analysis: Programming with R. / Chambers JM. – N.Y.: Springer-Verlag, 2008. – 514p.
- 19 Dethlefsen, Claus & Lundbye-Christensen Formulating State Space Models in R with Focus on Longitudinal Regression Models [Электронный ресурс] / Journal of Statistical Software – 2006. pp.1–15 – Режим доступа: https://www.researchgate.net/publication/26469366_Formulating_State_Space_Models_in_R_with_Focus_on_Longitudinal_Regression_Models.
- 20 Harvey A.C. Forecasting, Structural Time Series Models and the Kalman Filter / Harvey A.C. – Cambridge University Press , 1989.– 559 p.

- 21 Maximum likelihood // The free encyclopedia «Wikipedia» [электронный ресурс]. Режим доступа: http://en.wikipedia.org/wiki/Maximum_likelihood
- 22 Pradhan R.P. Forecasting Exchange Rate in India: An Application of Artificial Neural Network Model [Электронный ресурс] Pradhan R.P., Kumar R. // Journal of Mathematics Research. – 2010, Vol. 2, No. 4. – pp. 111 – 117. Режим доступа: <http://pdfs.semanticscholar.org/1fa7/2f9c2bb812692d28fbf09f7783c89bc7ac7a>
- 23 Prajakta S.K. Time series Forecasting using Holt-Winters Exponential Smoothing / Prajakta S.K. // Kanwal Rekhi School of Information Technology Journal [Электронный ресурс]. 2004. – 13p. режим доступа : http://www.it.iitb.ac.in/~praj/acads/seminar/04329008_ExponentialSmoothing.pdf
- 24 Francis J. Mulhern A nearest neighbor model for forecasting market response [Электронный ресурс] Francis J. Mulhern, Robert J. Caprara // International Journal of Forecasting –1994. – Vol. 10, No 2. – pp. 191–207 Режим доступа: <https://www.sciencedirect.com/science/article/pii/0169207094900027>.

ДОДАТОК А

ЛІСТИНГ ПРОГРАМИ

Файл Main.R

```
library("forecast")
library("tseries")
library("FKF")
library("MASS")
library("GMDH")
library("stats")

forecastDim<-1
data<-read.table("~/Forecasting/Data/DowJones.txt",header=TRUE, sep=",")

TSFull<-ts(data$DJ,start = c(1960), frequency =1)
TS<-ts(data$DJ[1:(length(data$DJ)-forecastDim)],start = c(1960), frequency =1)

df<-adf.test(TSFull, k=0)
df
plot(TS)
```

Файл Stats.R

```
U<-function(original , fitted)
{
  sqrt(sum((original-fitted)*(original-
fitted)))/(sqrt(sum(original*original))+sqrt(sum(fitted*fitted)))
}
MiAE<- function(residuals) { min(abs(residuals)) }

Rsquared<-function(original , fitted) {
  y_mean<-(sum(original)/length(original))
  1-sum((original-fitted)^2)/sum((original-y_mean)^2)
}

MPE <- function(residuals, original) { (sum(residuals/original)/length(residuals))*100 }
RMSE <- function(residuals) { sqrt(sum(residuals*residuals)/length(residuals)) }
MAE <- function(residuals){ max(abs(residuals)) }
ME <- function(residuals) { sum(residuals)/length(residuals) }
MAPE <- function (residuals, original) { sum(abs(residuals)/abs(original)/length(residuals))*100 }
```

Файл ExponentialSmoothing.R

```
HoltWintersSmoothing<-HoltWinters(TS, gamma=FALSE, l.start=TS[1],b.start=(TS[2]-TS[1]))
```

```

plot(HoltWintersSmoothing$fitted)

HoltWintersSmoothingForecast <- forecast.HoltWinters(HoltWintersSmoothing, h=forecastDim)
plot.forecast(HoltWintersSmoothingForecast)
plot(HoltWintersSmoothingForecast)

residuals<-(TS[3:52]-HoltWintersSmoothingForecast$fitted)

plotFunctionWithForecast(TSFull,c(0,0,HoltWintersSmoothing$fitted[1:(length(TSFull)-2-
forecastDim)]), HoltWintersSmoothingForecast$mean)

RMSE(residuals=residuals)
MAE(residuals=residuals)
Rsquared(original=TSFull[3:52] , fitted=HoltWintersSmoothingForecast$fitted)

originalForecast<-TSFull[(length(TSFull)-forecastDim+1):length(TSFull)]
forecastErrors<-originalForecast-as.numeric(HoltWintersSmoothingForecast$mean)

RMSE(residuals=forecastErrors)
ME(residuals=forecastErrors)
MPE(residuals = forecastErrors, original = originalForecast)
MAPE(residuals = forecastErrors, original = originalForecast)
MAE(residuals=forecastErrors)
MiAE(residuals=forecastErrors)
U(original=TSFull[(length(TSFull)-forecastDim+1): length(TSFull)] ,
fitted=HoltWintersSmoothingForecast$forecasts)

```

Файл GMDH.R

```

studyDim<- as.integer(length(TS)*0.7)

out <- fcast(TS, input = studyDim, layer =2 , f.number = forecastDim, plotit=TRUE)

out$forecasts
out$fitted
out$MSE
residuals<-(TS[studyDim+1:(length(TS)- studyDim)]-out$fitted)
plot((1:length(residuals)),residuals, type="b", xlab="Time" )

RMSE(residuals=residuals)
MAE(residuals=residuals)
Rsquared(original=TSFull[(studyDim+1):(length(TSFull)-forecastDim)] , fitted=out$fitted)

plotFunctionWithForecast(as.numeric(TSFull), c(vector(mode="numeric",
length=studyDim),out$fitted[1:(length(out$fitted)]), out$forecasts )

originalForecast<-TSFull[(length(TSFull)-forecastDim+1):length(TSFull)]

```

```

forecastErrors<-originalForecast-as.numeric(out$forecasts)
RMSE(residuals=forecastErrors)
ME(residuals=forecastErrors)
MPE(residuals = forecastErrors, original = originalForecast)
MAPE(residuals = forecastErrors, original = originalForecast)
MAE(residuals=forecastErrors)
MiAE(residuals=forecastErrors)
U(original=TSFull[(length(TSFull)-forecastDim+1): length(TSFull)] , fitted=out$forecasts)

```

Файл KalmanFilter.R

```

acf(TSFull, main="TS ACF")
pacf(TSFull, main="TS PACF")

plot(TS, main="Time series")
arma<-auto.arima(TS, trace=TRUE)
arma

armafr<-forecast.Arima(arma, h=forecastDim)

acf(armafr$residuals)
plot(armafr$residuals)

Box.test(armafr$residuals, lag=10, type="Ljung-Box")

plot(armafr$residuals,ylab="залишки моделі")
plotForecastErrors(armafr$residuals)
plotFunctionWithForecast(as.numeric(TSFull), armafr$fitted,armafr$mean )

RMSE(residuals=armafr$residuals)
MAE(residuals=armafr$residuals)
Rsquared(original=TSFull[1:(length(TSFull)-forecastDim)] , fitted=armafr$fitted)
forecastErrors<-TSFull[(length(TSFull)-forecastDim+1): length(TSFull)]-armafr$mean
forecastErrors

RMSE(residuals=forecastErrors)
ME(residuals=forecastErrors)
MPE(residuals = forecastErrors, original = armafr$mean)
MAPE(residuals = forecastErrors, original = armafr$mean)
MAE(residuals=forecastErrors)
MiAE(residuals=forecastErrors)
U(original=TSFull[(length(TSFull)-forecastDim+1): length(TSFull)] , fitted=armafr$mean)

```

Файл NNforecast.R

```

NNforecast<- function(dim)
{
dim<-1

```

```

dt<-read.table("~/Forecasting/Data/NNDowJonesFORECAST.txt",header=TRUE, sep=",")
NNfrc<-ts(dt$forecast,start = c(2000), frequency =1)
plotFunctionWithForecast(as.numeric(TSFull), NNfrc[1:(length(TSFull)-dim)],NNfrc[(length(TSFull)-
dim+1):length(TSFull)] )

forecastErrors<-TSFull[(length(TSFull)-dim+1):length(TSFull)]-NNfrc[(length(TSFull)-
dim+1):length(TSFull)]
forecastErrors

RMSE(residuals=forecastErrors)
ME(residuals=forecastErrors)
MPE(residuals = forecastErrors, original = armafr$mean)
MAPE(residuals = forecastErrors, original = armafr$mean)
MAE(residuals=forecastErrors)
MiAE(residuals=forecastErrors)
U(original=TSFull[(length(TSFull)-dim+1): length(TSFull)] , fitted=NNfrc[(length(TSFull)-
dim+1):length(TSFull)])
}

```

Файл PlotFunctionWithForecast.R

```

plotFunctionWithForecast= function(tsFull, ts, forecast){

  aproximated<-ts
  real<-tsFull

  miny<-min(c(real,aproximated,forecast))
  miny
  maxy<-max(c(real,aproximated,forecast))
  maxy
  plot((1:length(real)),
       real, type="l",col="black",
       ylab="",xlab="time",
       ylim=c(miny,maxy),
       xlim=c(1,length(real)), lwd=2)

  par(new=TRUE)

  plot((1:length(aproximated)

```

ДОДАТОК Б

ЧАСОВІ РЯДИ

Ряд з цінами ф'ючесів на золото

date	price
01.03.2019	11 154,35
01.04.2019	11 246,54
01.05.2019	11 247,69
01.06.2019	11 237,62
01.09.2019	11 169,79
01.10.2019	11 183,33
01.11.2019	11 226,78
01.12.2019	11 204,15
01.13.2019	11 227,17
01.17.2019	11 196,29
01.18.2019	11 196,10
01.19.2019	11 148,85
01.20.2019	11 192,79
01.21.2019	11 170,62
01 .24.2019	11 249,28
01.25.2019	11 339,05
01.26.2019	11 313,12
01.27.2019	11 283,19
01.30.2019	11 205,24
01.31.2019	11 222,95
02.01.2019	11 207,24
02.02.2019	11 215,38
02.03.2019	11 310,74
02.06.2019	11 264,11
02.07.2019	11 243,38
02.08.2019	11 251,80
02.09.2019	11 327,68
02.10.2019	11 377,72
02.13.2019	11 433,93
02.14.2019	11 468,87
02.15.2019	11 514,40
02.16.2019	11 504,19

02.17.2019	11 510,91
02.21.2019	11 578,25
02.22.2019	11 550,18
02.23.2019	11 556,15
02.24.2019	11 541,29
02.27.2019	11 558,35
02.28.2019	11 512,39
03.01.2019	11 661,22
03.02.2019	11 575,91
03.03.2019	11 598,37
03.06.2019	11 547,79
03.07.2019	11 506,32
03.08.2019	11 448,21
03.09.2019	11 457,64
03.10.2019	11 500,76
03.13.2019	11 519,44
03.14.2019	11 459,29
03.15.2019	11 590,31
03.16.2019	11 600,23
03.17.2019	11 589,01
03.20.2019	11 556,93
03.21.2019	11 420,43
03.22.2019	11 426,69
03.23.2019	11 430,89
03.24.2019	11 418,89
03.27.2019	11 414,33
03.28.2019	11 493,84
03.29.2019	11 496,67
03.30.2019	11 519,84
03.31.2019	11 492,85
04.03.2019	11 488,15
04.04.2019	11 478,25
04.05.2019	11 444,99

04.06.2019	11 471,75
04.07.2019	11 466,87
04.10.2019	11 475,13
04.11.2019	11 462,07
04.12.2019	11 427,01

04.13.2019	11 380,80
04.14.2019	11 420,85
04.17.2019	11 452,85
04.18.2019	11 330,42

Ряд EUR_USD

data	EUR_USD
01.12.2018	1.0662
02.12.2018	1.0673
04.12.2018	1.0567
05.12.2018	1.0764
06.12.2018	1.0719
07.12.2018	1.0753
08.12.2018	1.0615
09.12.2018	1.0562
11.12.2018	1.0547
12.12.2018	1.0635
13.12.2018	1.0627
14.12.2018	1.0536
15.12.2018	1.0414
16.12.2018	1.0451
18.12.2018	1.0447
19.12.2018	1.0402
20.12.2018	1.0387
21.12.2018	1.0425
22.12.2018	1.0437
23.12.2018	1.0456
25.12.2018	1.0454
26.12.2018	1.0456
27.12.2018	1.0457
28.12.2018	1.0412
29.12.2018	1.0492
30.12.2018	1.0516
01.01.2019	1.0526
02.01.2019	1.0459

03.01.2019	1.0406
04.01.2019	1.0488
05.01.2019	1.0607
06.01.2019	1.0533
08.01.2019	1.0533
09.01.2019	1.0574
10.01.2019	1.0554
11.01.2019	1.0582
12.01.2019	1.0613
13.01.2019	1.0644
15.01.2019	1.0627
16.01.2019	1.0603
17.01.2019	1.0714
18.01.2019	1.0631
19.01.2019	1.0664
20.01.2019	1.0703
22.01.2019	1.0713
23.01.2019	1.0765
24.01.2019	1.0731
25.01.2019	1.0748
26.01.2019	1.0682
27.01.2019	1.0699
30.01.2019	1.0695
31.01.2019	1.0798
01.02.2019	1.0769
02.02.2019	1.0758
03.02.2019	1.0784
06.02.2019	1.0784
07.02.2019	1.0681

08.02.2019	1.0681
09.02.2019	1.0656
10.02.2019	1.0641
13.02.2019	1.0598
14.02.2019	1.0577
15.02.2019	1.0601
16.02.2019	1.0674
17.02.2019	1.0614
20.02.2019	1.0614
21.02.2019	1.0537
22.02.2019	1.0556
23.02.2019	1.0581
24.02.2019	1.0563
27.02.2019	1.0588
28.02.2019	1.0577
01.03.2019	1.0548
02.03.2019	1.0507
03.03.2019	1.0623
06.03.2019	1.0581
07.03.2019	1.0566
08.03.2019	1.0541
09.03.2019	1.0577
10.03.2019	1.0672
13.03.2019	1.0654
14.03.2019	1.0604

15.03.2019	1.0735
16.03.2019	1.0766
17.03.2019	1.0738
20.03.2019	1.0739
21.03.2019	1.0811
22.03.2019	1.0797
23.03.2019	1.0783
24.03.2019	1.0799
27.03.2019	1.0865
28.03.2019	1.0813
29.03.2019	1.0767
30.03.2019	1.0675
31.03.2019	1.0652
03.04.2019	1.0671
04.04.2019	1.0674
05.04.2019	1.0663
06.04.2019	1.0644
07.04.2019	1.0644
10.04.2019	1.0596
11.04.2019	1.0602
12.04.2019	1.0666
13.04.2019	1.0614
14.04.2019	1.0612
17.04.2019	1.0643
18.04.2019	1.0734

Таблиця Б.1 – Індекси для побудови мережі Байєса

Data	DJ	S&P500	Nasdaq	Nasdaq_Com	NYSE_Com
1/2/2017	17832.99	2058.2	4230.24	4726.81	10830.92
1/5/2017	17501.65	2020.58	4160.96	4652.57	10607.68
1/6/2017	17371.64	2002.61	4110.83	4592.74	10514.87
1/7/2017	17584.52	2025.9	4160	4650.47	10627.13
1/8/2017	17907.87	2062.14	4240.55	4736.19	10800.55
1/9/2017	17737.37	2044.81	4213.28	4704.07	10711.41
1/12/2017	17640.84	2028.26	4169.97	4664.71	10640.74
1/13/2017	17613.68	2023.03	4166.2	4661.5	10622.96
1/14/2017	17427.09	2011.27	4145.84	4639.32	10565.41
1/15/2017	17320.71	1992.67	4089.65	4570.82	10514.6
1/16/2017	17511.57	2019.42	4142.14	4634.38	10660.32
1/20/2017	17515.23	2022.55	4171.21	4654.85	10658.62
1/21/2017	17554.28	2032.12	4192.09	4667.42	10720.4
1/22/2017	17813.98	2063.15	4270.36	4750.4	10859.34
1/23/2017	17672.6	2051.82	4278.14	4757.88	10788.33
1/26/2017	17678.7	2057.09	4275.72	4771.76	10847.15
1/27/2017	17387.21	2029.55	4165.5	4681.5	10781.83
1/28/2017	17191.37	2002.16	4140.38	4637.99	10603.86
1/29/2017	17416.85	2021.25	4181.35	4683.41	10691.63
1/30/2017	17164.95	1994.99	4148.43	4635.24	10537.22
2/2/2017	17361.04	2020.85	4188.59	4676.69	10669.03
2/3/2017	17666.4	2050.03	4229.15	4727.74	10847.41
2/4/2017	17673.02	2041.51	4221.2	4716.7	10775.04
2/5/2017	17884.88	2062.52	4256.18	4765.1	10896.06
2/6/2017	17824.29	2055.47	4228.68	4744.4	10847.51
2/9/2017	17729.21	2046.74	4216.09	4726.01	10826.59
2/10/2017	17868.76	2068.59	4281.15	4787.65	10915.14
2/11/2017	17862.14	2068.53	4297.28	4801.18	10889.05
2/12/2017	17972.38	2088.48	4347.97	4857.61	10995.31
2/13/2017	18019.35	2096.99	4384.03	4893.84	11042.69
2/17/2017	18047.58	2100.34	4385.34	4899.27	11054.46
2/18/2017	18029.85	2099.68	4390.91	4906.36	11064.09
2/19/2017	17985.77	2097.45	4411.86	4924.7	11038.26
2/20/2017	18140.44	2110.3	4443.05	4955.97	11108.67
2/23/2017	18116.84	2109.66	4449.49	4960.97	11074.58
2/24/2017	18209.19	2115.48	4451.03	4968.12	11122.11
2/25/2017	18224.57	2113.86	4440.59	4967.14	11117.93
2/26/2017	18214.42	2110.74	4462.27	4987.89	11082.1
2/27/2017	18132.7	2104.5	4440.67	4963.53	11062.79
3/2/2017	18288.63	2117.39	4483.05	5008.1	11102.82
3/3/2017	18203.37	2107.78	4458.78	4979.9	11051.22
3/4/2017	18096.9	2098.53	4445.36	4967.14	11004.68

Продовження таблиці Б.1

3/5/2017	18135.72	2101.04	4452.06	4982.81	11014.26
3/6/2017	17856.78	2071.26	4399.23	4927.37	10842.17
3/9/2017	17995.72	2079.43	4412.99	4942.44	10866.62
3/10/2017	17662.94	2044.16	4329.08	4859.8	10678.41
3/11/2017	17635.39	2040.24	4305.38	4849.94	10678.76
3/12/2017	17895.22	2065.95	4336.23	4893.29	10820.87
3/13/2017	17749.31	2053.4	4314.9	4871.76	10751.02
3/16/2017	17977.42	2081.19	4370.47	4929.51	10881.18
3/17/2017	17849.08	2074.28	4375.62	4937.44	10862.43
3/18/2017	18076.19	2099.5	4422.5	4982.83	11019.72
3/19/2017	17959.03	2089.27	4426.82	4992.38	10931.36
3/20/2017	18127.65	2108.1	4458.54	5026.42	11070.53
3/23/2017	18116.04	2104.42	4445.54	5010.97	11070.89
3/24/2017	18011.14	2091.5	4430.98	4994.73	11019.75
3/25/2017	17718.54	2061.05	4329.29	4876.52	10897.51
3/26/2017	17678.23	2056.15	4315.26	4863.36	10865.15
3/27/2017	17712.66	2061.02	4332.86	4891.22	10875.14
3/30/2017	17976.31	2086.24	4382.82	4947.44	10989.16
3/31/2017	17776.12	2067.89	4333.69	4900.88	10899.18
4/1/2017	17698.18	2059.69	4311.26	4880.23	10891.61
4/2/2017	17763.24	2066.96	4316.01	4886.94	10953.16
4/6/2017	17880.85	2080.62	4350.98	4917.32	11033.12
4/7/2017	17875.42	2076.33	4344.08	4910.23	11013.24
4/8/2017	17902.51	2081.9	4375.96	4950.82	11032.99
4/9/2017	17958.73	2091.18	4403.95	4974.56	11065.39
4/10/2017	18057.65	2102.06	4422.02	4995.98	11112.68
4/13/2017	17977.04	2092.43	4409.21	4988.25	11056.76
4/14/2017	18036.7	2095.84	4397.65	4977.29	11108.07
4/15/2017	18112.61	2106.63	4424.53	5011.02	11171.06
4/16/2017	18105.77	2104.99	4418.83	5007.79	11169.75
4/17/2017	17826.3	2081.18	4351.8	4931.81	11058.45
4/20/2017	18034.93	2100.4	4417.4	4994.6	11116.57
4/21/2017	17949.59	2097.29	4435.98	5014.1	11100.38
4/22/2017	18038.27	2107.96	4460.81	5035.17	11144.45
4/23/2017	18058.69	2112.93	4477.07	5056.06	11191.48
4/24/2017	18080.14	2117.69	4536.78	5092.09	11192.93
4/27/2017	18037.97	2108.92	4525.03	5060.25	11167.31
4/28/2017	18110.14	2114.76	4515.27	5055.42	11203.08
4/29/2017	18035.53	2106.85	4488.15	5023.64	11154.11
4/30/2017	17840.52	2085.51	4414.25	4941.42	11049.74
5/1/2017	18024.06	2108.29	4479.06	5005.39	11140.36
5/4/2017	18070.4	2114.49	4482.61	5016.93	11173.28
5/5/2017	17928.2	2089.46	4409.39	4939.33	11049.34
5/6/2017	17841.98	2080.15	4379.75	4919.64	11015.99

Продовження таблиці Б.1

5/7/2017	17924.06	2088	4401.55	4945.54	11036.4
5/8/2017	18191.11	2116.1	4458.65	5003.55	11196.5
5/11/2017	18105.17	2105.33	4438.64	4993.57	11140.74
5/12/2017	18068.23	2099.12	4420.65	4976.19	11113.17
5/13/2017	18060.49	2098.48	4426.56	4981.69	11117
5/14/2017	18252.24	2121.1	4495.04	5050.8	11207.33
5/15/2017	18272.56	2122.73	4494.29	5048.29	11228.35
5/18/2017	18298.88	2129.2	4512.95	5078.44	11227.89
5/19/2017	18312.39	2127.83	4503.72	5070.03	11213.15
5/20/2017	18285.4	2125.85	4505.3	5071.74	11210.24
5/21/2017	18285.74	2130.82	4529.47	5090.79	11239.66
5/22/2017	18232.02	2126.06	4527.16	5089.36	11197.69
5/26/2017	18041.54	2104.2	4472.96	5032.75	11065.32
5/27/2017	18162.99	2123.48	4546.06	5106.59	11143.35
5/28/2017	18126.12	2120.79	4535.66	5097.98	11132.74
5/29/2017	18010.68	2107.39	4508.25	5070.02	11056.3
6/1/2017	18040.37	2111.73	4521.85	5082.93	11061.68
6/2/2017	18011.94	2109.6	4508.89	5076.52	11080.89
6/3/2017	18076.27	2114.07	4519.86	5099.23	11107.93
6/4/2017	17905.58	2095.84	4487.03	5059.12	11000.39
6/5/2017	17849.46	2092.83	4477.19	5068.46	10979.33
6/8/2017	17766.55	2079.28	4431.52	5021.63	10917.82
6/9/2017	17764.04	2080.15	4427.61	5013.86	10915.64
6/10/2017	18000.4	2105.2	4485.46	5076.69	11057.99
6/11/2017	18039.37	2108.86	4488.47	5082.51	11080.51
6/12/2017	17898.84	2094.11	4453.79	5051.1	11009.91
6/15/2017	17791.17	2084.43	4432.92	5029.97	10949.54
6/16/2017	17904.48	2096.29	4455.9	5055.55	10997.66
6/17/2017	17935.74	2100.44	4468.98	5064.88	11013.66
6/18/2017	18115.84	2121.24	4531.85	5132.95	11100.47
6/19/2017	18015.95	2109.99	4513.42	5117	11038.96
6/22/2017	18119.78	2122.85	4544.04	5153.97	11129.61
6/23/2017	18144.07	2124.2	4548.74	5160.1	11149.36
6/24/2017	17966.07	2108.58	4524.6	5122.41	11066.14
6/25/2017	17890.36	2102.31	4514.01	5112.19	11035.73
6/26/2017	17946.68	2101.49	4484.28	5080.51	11040.31
6/29/2017	17596.35	2057.64	4379.86	4958.47	10790.33
6/30/2017	17619.51	2063.11	4396.76	4986.87	10805.2
7/1/2017	17757.91	2077.42	4429.23	5013.12	10858.99
7/2/2017	17730.11	2076.78	4433.39	5009.21	10864.82
7/6/2017	17683.58	2068.76	4418.66	4991.94	10782.68
7/7/2017	17776.91	2081.34	4429.06	4997.46	10829.32
7/8/2017	17515.42	2046.68	4351.58	4909.76	10636.8
7/9/2017	17548.62	2051.31	4352.03	4922.4	10689.51

Продовження таблиці Б.1

7/10/2017	17760.41	2076.62	4420.15	4997.7	10853.92
7/13/2017	17977.68	2099.6	4494.47	5071.51	10944.85
7/14/2017	18053.58	2108.95	4523.47	5104.89	10993.26
7/15/2017	18050.17	2107.4	4528.36	5098.94	10958.83
7/16/2017	18120.25	2124.29	4594.9	5163.18	11024.87
7/17/2017	18086.45	2126.64	4661.6	5210.14	10987.17
7/20/2017	18100.41	2128.28	4679.68	5218.86	10969.49
7/21/2017	17919.29	2119.21	4674.4	5208.12	10919
7/22/2017	17851.04	2114.15	4623.24	5171.77	10899.86
7/23/2017	17731.92	2102.15	4602.83	5146.41	10840.73
7/24/2017	17568.53	2079.65	4557.38	5088.63	10721.95
7/27/2017	17440.59	2067.64	4519.09	5039.78	10656.61
7/28/2017	17630.27	2093.25	4560.23	5089.21	10790.18
7/29/2017	17751.39	2108.57	4581.6	5111.73	10883.91
7/30/2017	17745.98	2108.63	4598.59	5128.78	10881.75
7/31/2017	17689.86	2103.84	4588.91	5128.28	10882.28
8/3/2017	17598.2	2098.04	4580.46	5115.38	10841.26
8/4/2017	17550.69	2093.32	4567.47	5105.55	10824.7
8/5/2017	17540.47	2099.84	4602.24	5139.94	10847.85
8/6/2017	17419.75	2083.56	4528.51	5056.44	10799.44
8/7/2017	17373.38	2077.57	4520.08	5043.54	10763.15
8/10/2017	17615.17	2104.18	4573.11	5101.8	10906.03
8/11/2017	17402.84	2084.07	4513.99	5036.79	10802.09
8/12/2017	17402.51	2086.05	4528.19	5044.39	10779.46
8/13/2017	17408.25	2083.39	4519.32	5033.56	10745.31
8/14/2017	17477.4	2091.54	4530.74	5048.23	10782.24
8/17/2017	17545.18	2102.44	4566.37	5091.7	10818.57
8/18/2017	17511.34	2096.92	4542.07	5059.35	10792.8
8/19/2017	17348.73	2079.61	4510.9	5019.05	10687.03
8/20/2017	16990.69	2035.73	4385.13	4877.49	10475.75
8/21/2017	16459.75	1970.89	4197.27	4706.04	10195.69
8/24/2017	15871.35	1893.21	4038.6	4526.25	9789.94
8/25/2017	15666.44	1867.61	4016.32	4506.49	9692.39
8/26/2017	16285.51	1940.51	4219.6	4697.54	9979.67
8/27/2017	16654.77	1987.66	4324.82	4812.71	10230.51
8/28/2017	16643.01	1988.87	4329.12	4828.32	10242.06
8/31/2017	16528.03	1972.18	4274.58	4776.51	10176.5
9/1/2017	16058.35	1913.85	4142.63	4636.1	9874.67
9/2/2017	16351.38	1948.86	4256.23	4749.98	10013.36
9/3/2017	16374.76	1951.13	4234.65	4733.5	10049.29
9/4/2017	16102.38	1921.22	4184.72	4683.92	9871.86
9/8/2017	16492.68	1969.41	4303.53	4811.93	10109.68
9/9/2017	16253.57	1942.04	4253.45	4756.53	9988.91
9/10/2017	16330.4	1952.29	4296.3	4796.25	10019.39

Продовження таблиці Б.1

9/11/2017	16433.09	1961.05	4323.23	4822.34	10040.22
9/14/2017	16370.96	1953.03	4308.76	4805.76	9988.59
9/15/2017	16599.85	1978.09	4360.28	4860.52	10105.87
9/16/2017	16739.95	1995.31	4383.67	4889.24	10236.72
9/17/2017	16674.74	1990.2	4384.96	4893.95	10215.54
9/18/2017	16384.58	1958.03	4323.86	4827.23	10031.6
9/21/2017	16510.19	1966.97	4336.13	4828.95	10066.07
9/22/2017	16330.47	1942.74	4274.1	4756.72	9912.6
9/23/2017	16279.89	1938.76	4273.53	4752.74	9867.9
9/24/2017	16201.32	1932.24	4261.1	4734.48	9837.81
9/25/2017	16314.67	1931.34	4224.74	4686.5	9857.25
9/28/2017	16001.89	1881.77	4103.47	4543.97	9601.42
9/29/2017	16049.13	1884.09	4083.37	4517.32	9622.07
9/30/2017	16284.7	1920.03	4181.06	4620.16	9799.69
10/1/2017	16272.01	1923.82	4192.96	4627.08	9820.95
10/2/2017	16472.37	1951.36	4267.45	4707.77	9973.56
10/5/2017	16776.43	1987.05	4331.43	4781.26	10167.93
10/6/2017	16790.19	1979.92	4306.7	4748.36	10156.38
10/7/2017	16912.29	1995.83	4333.46	4791.15	10258.52
10/8/2017	17050.75	2013.43	4350.15	4810.79	10361.04
10/9/2017	17084.49	2014.89	4370.01	4830.47	10361.26
10/12/2017	17131.86	2017.46	4382.56	4838.64	10356.49
10/13/2017	17081.89	2003.69	4352.94	4796.61	10262.94
10/14/2017	16924.75	1994.24	4346.92	4782.85	10229.67
10/15/2017	17141.75	2023.86	4418.22	4870.1	10378.27
10/16/2017	17215.97	2033.11	4438.62	4886.69	10421.91
10/19/2017	17230.54	2033.66	4462.07	4905.47	10391.6
10/20/2017	17217.11	2030.77	4438.18	4880.97	10383.42
10/21/2017	17168.61	2018.94	4413.52	4840.12	10306.44
10/22/2017	17489.16	2052.51	4503.22	4920.05	10448.83
10/23/2017	17646.7	2075.15	4624.09	5031.86	10506.51
10/26/2017	17623.05	2071.18	4630.85	5034.7	10464.62
10/27/2017	17581.43	2065.89	4639.23	5030.15	10400.18
10/28/2017	17779.52	2090.35	4678.57	5095.69	10538.2
10/29/2017	17755.8	2089.41	4670.77	5074.27	10498.16
10/30/2017	17663.54	2079.36	4648.83	5053.75	10460.96
11/2/2017	17828.76	2104.05	4703.91	5127.15	10579.87
11/3/2017	17918.15	2109.79	4719.05	5145.13	10609.92
11/4/2017	17867.58	2102.31	4717.55	5142.48	10563.96
11/5/2017	17863.43	2099.93	4702.9	5127.74	10548.36
11/6/2017	17910.33	2099.2	4707.23	5147.12	10513.36
11/9/2017	17730.48	2078.58	4655.52	5095.3	10415.19
11/10/2017	17758.21	2081.72	4641.73	5083.24	10432.35
11/11/2017	17702.22	2075	4636.94	5067.02	10399.9

Продовження таблиці Б.1

11/12/2017	17448.07	2045.97	4588.93	5005.08	10235.95
11/13/2017	17245.24	2023.04	4502.14	4927.88	10155.07
11/16/2017	17483.01	2053.19	4565.47	4984.62	10299.4
11/17/2017	17489.5	2050.44	4565.82	4986.02	10285.76
11/18/2017	17737.16	2083.58	4653.44	5075.2	10429.6
11/19/2017	17732.75	2081.24	4655.35	5073.64	10431.46
11/20/2017	17823.81	2089.17	4686.36	5104.92	10444.2
11/23/2017	17792.68	2086.59	4674.79	5102.48	10421.41
11/24/2017	17812.19	2089.14	4669.41	5102.81	10449.72
11/25/2017	17813.39	2088.87	4672.88	5116.14	10450.36
11/27/2017	17798.49	2090.11	4680.47	5127.52	10450.53
11/30/2017	17719.92	2080.41	4664.51	5108.67	10409.58
12/1/2017	17888.35	2102.63	4716.16	5156.31	10519.59
12/2/2017	17729.68	2079.51	4686.22	5123.22	10389.2
12/3/2017	17477.67	2049.62	4606.76	5037.53	10257.38
12/4/2017	17847.63	2091.69	4716.12	5142.27	10408.86
12/7/2017	17730.51	2077.07	4693.62	5101.81	10304.53
12/8/2017	17568	2063.59	4694.77	5098.24	10198.37
12/9/2017	17492.3	2047.62	4622.14	5022.86	10162.36
12/10/2017	17574.75	2052.23	4643.05	5045.17	10172.04
12/11/2017	17265.21	2012.37	4537.56	4933.47	9976.65
12/14/2017	17368.5	2021.94	4570.16	4952.23	9986.76
12/15/2017	17524.91	2043.41	4597.33	4995.36	10110.94
12/16/2017	17749.09	2073.07	4664.56	5071.13	10267.83
12/17/2017	17495.84	2041.89	4598.14	5002.55	10112.63
12/18/2017	17128.55	2005.55	4514.82	4923.08	9967.64
12/21/2017	17251.62	2021.15	4560.55	4968.92	10023.03
12/22/2017	17417.27	2038.97	4589.26	5001.11	10112.4
12/23/2017	17602.61	2064.29	4624.59	5045.93	10276.28
12/24/2017	17552.17	2060.99	4622.77	5048.49	10258.55
12/28/2017	17528.27	2056.5	4621.35	5040.98	10217.4
12/29/2017	17720.98	2078.36	4691.08	5107.94	10306.6
12/30/2017	17603.87	2063.36	4652.01	5065.85	10225.83
12/31/2017	17425.03	2043.94	4593.27	5007.41	10143.42
1/4/2018	17148.94	2012.66	4497.86	4903.09	10001.56
1/5/2018	17158.66	2016.71	4484.18	4891.43	10028.06
1/6/2018	16906.51	1990.26	4443.98	4835.77	9868.26
1/7/2018	16514.1	1943.09	4305.72	4689.43	9650.42
1/8/2018	16346.45	1922.03	4270.78	4643.63	9528.77
1/11/2018	16398.57	1923.67	4283.55	4637.99	9527.66
1/12/2018	16516.22	1938.68	4333.72	4685.92	9576.52
1/13/2018	16151.41	1890.28	4183.12	4526.06	9374.23
1/14/2018	16379.05	1921.84	4272.97	4615	9516.55
1/15/2018	15988.08	1880.33	4141.08	4488.42	9299.62

Продовження таблиці Б.1

1/19/2018	16016.02	1881.33	4147.07	4476.95	9301.74
1/20/2018	15766.74	1859.33	4136.28	4471.69	9156.84
1/21/2018	15882.68	1868.99	4142.6	4472.06	9226.57
1/22/2018	16093.51	1906.9	4259.77	4591.18	9426.91
1/25/2018	15885.22	1877.08	4196.53	4518.49	9262.99
1/26/2018	16167.23	1903.63	4233.85	4567.67	9429.11
1/27/2018	15944.46	1882.95	4128.86	4468.17	9373.78
1/28/2018	16069.64	1893.36	4186.06	4506.68	9397.79
1/29/2018	16466.3	1940.24	4279.17	4613.95	9632.7
2/1/2018	16449.18	1939.38	4286.56	4620.37	9616.69
2/2/2018	16153.54	1903.03	4193.1	4516.95	9413.78
2/3/2018	16336.66	1912.53	4171.97	4504.24	9495.21
2/4/2018	16416.58	1915.45	4167.77	4509.56	9534.3
2/5/2018	16204.97	1880.05	4024.47	4363.14	9390.33
2/8/2018	16027.05	1853.44	3960.67	4283.75	9231.33
2/9/2018	16014.38	1852.21	3947.8	4268.76	9187.79
2/10/2018	15914.74	1851.86	3966.28	4283.59	9176.72
2/11/2018	15660.18	1829.08	3962.22	4266.84	9029.88
2/12/2018	15973.84	1864.78	4019.19	4337.51	9229.68
2/16/2018	16196.41	1895.58	4104.33	4435.95	9375.39
2/17/2018	16453.83	1926.82	4199.1	4534.06	9532.28
2/18/2018	16413.43	1917.83	4151.49	4487.54	9504.98
2/19/2018	16391.99	1917.78	4164.09	4504.43	9485.96
2/22/2018	16620.66	1945.5	4231.27	4570.61	9616.32
2/23/2018	16431.78	1921.27	4162.54	4503.58	9494.32
2/24/2018	16484.99	1929.8	4200.66	4542.61	9506.07
2/25/2018	16697.29	1951.7	4241.06	4582.2	9625.27
2/26/2018	16639.97	1948.05	4235.62	4590.47	9619.79
2/29/2018	16516.5	1932.23	4201.12	4557.95	9559.53
3/1/2018	16865.08	1978.35	4333.61	4689.6	9771
3/2/2018	16899.32	1986.45	4334.39	4703.42	9837.59
3/3/2018	16943.9	1993.4	4325.86	4707.42	9907.97
3/4/2018	17006.77	1999.99	4329	4717.02	9968.41
3/7/2018	17073.95	2001.76	4303.33	4708.25	10006.19
3/8/2018	16964.1	1979.26	4265.51	4648.82	9868.52
3/9/2018	17000.36	1989.26	4293.23	4674.38	9924.81
3/10/2018	16995.13	1989.57	4286.97	4662.16	9918.07
3/11/2018	17213.31	2022.19	4361.83	4748.47	10104.19
3/14/2018	17229.13	2019.64	4367.33	4750.28	10074.27
3/15/2018	17251.53	2015.93	4367.08	4728.67	10013.6
3/16/2018	17325.76	2027.22	4404.2	4763.97	10084.17
3/17/2018	17481.49	2040.59	4400.26	4774.98	10190.4
3/18/2018	17602.3	2049.58	4410.83	4795.65	10223.43
3/21/2018	17623.87	2051.6	4426.98	4808.87	10219.69

Продовження таблиці Б.1

3/22/2018	17582.57	2049.8	4437.62	4821.66	10194.2
3/23/2018	17502.59	2036.71	4402.56	4768.86	10099.68
3/24/2018	17515.73	2035.94	4405.53	4773.51	10086.6
3/28/2018	17535.39	2037.05	4398.07	4766.79	10106.44
3/29/2018	17633.11	2055.01	4467.72	4846.62	10190.41
3/30/2018	17716.66	2063.95	4490.88	4869.29	10236.95
3/31/2018	17685.09	2059.74	4483.65	4869.85	10207.38
4/1/2018	17792.75	2072.78	4532.08	4914.54	10219.96
4/4/2018	17737	2066.13	4511.7	4891.8	10183
4/5/2018	17603.32	2045.17	4470.75	4843.93	10061.08
4/6/2018	17716.05	2066.66	4543.78	4920.72	10159.85
4/7/2018	17541.96	2041.91	4475.28	4848.37	10045.15
4/8/2018	17576.96	2047.6	4474.93	4850.69	10119.69
4/11/2018	17556.41	2041.99	4458.7	4833.4	10114.58
4/12/2018	17721.25	2061.72	4496.05	4872.09	10238.69
4/13/2018	17908.28	2082.42	4554.72	4947.42	10359.51
4/14/2018	17926.43	2082.78	4554.62	4945.89	10360.24
4/15/2018	17897.46	2080.73	4543.16	4938.22	10355.57
4/18/2018	18004.16	2094.34	4565.42	4960.02	10424.93
4/19/2018	18053.6	2100.8	4537.09	4940.33	10517.3
4/20/2018	18096.27	2102.4	4540.44	4948.13	10527.79
4/21/2018	17982.52	2091.48	4540.8	4945.89	10454.82
4/22/2018	18003.75	2091.58	4474.19	4906.23	10511
4/25/2018	17977.24	2087.79	4474.13	4895.79	10465.65
4/26/2018	17990.32	2091.7	4452.97	4888.28	10522.58
4/27/2018	18041.55	2095.15	4416.64	4863.14	10571.83
4/28/2018	17830.76	2075.81	4363.99	4805.29	10480.45
4/29/2018	17773.64	2065.3	4341.3	4775.36	10436.92
5/2/2018	17891.16	2081.43	4381.3	4817.59	10496.33
5/3/2018	17750.91	2063.37	4341.39	4763.22	10365.83
5/4/2018	17651.26	2051.12	4312.01	4725.64	10286.24
5/5/2018	17660.71	2050.63	4309.48	4717.09	10281.32
5/6/2018	17740.63	2057.14	4330.17	4736.15	10308.83
5/9/2018	17705.91	2058.69	4341.23	4750.21	10291.59
5/10/2018	17928.35	2084.39	4402.04	4809.88	10424.11
5/11/2018	17711.12	2064.46	4361.03	4760.69	10332.24
5/12/2018	17720.5	2064.11	4342.81	4737.33	10334.38
5/13/2018	17535.32	2046.61	4326.53	4717.68	10228.06
5/16/2018	17710.71	2066.66	4380.48	4775.46	10332.47
5/17/2018	17529.98	2047.21	4322.56	4715.73	10257.61
5/18/2018	17526.62	2047.63	4338.23	4739.12	10239.65
5/19/2018	17435.4	2040.04	4315.58	4712.53	10192.48
5/20/2018	17500.94	2052.32	4362.9	4769.56	10250.49
5/23/2018	17492.93	2048.04	4355.1	4765.78	10226.56

Продовження таблиці Б.1

5/24/2018	17706.05	2076.06	4444.3	4861.06	10353.56
5/25/2018	17851.51	2090.54	4476.43	4894.89	10446.58
5/26/2018	17828.29	2090.1	4487.96	4901.77	10439.61
5/27/2018	17873.22	2099.06	4512.54	4933.51	10469.52
5/31/2018	17787.2	2096.96	4523.89	4948.06	10441
6/1/2018	17789.67	2099.33	4520.07	4952.25	10454.93
6/2/2018	17838.56	2105.26	4531.35	4971.36	10487.58
6/3/2018	17807.06	2099.13	4509.79	4942.52	10487.94
6/6/2018	17920.33	2109.41	4523.63	4968.71	10554.82
6/7/2018	17938.28	2112.13	4513.09	4961.75	10599.22
6/8/2018	18005.05	2119.12	4520.63	4974.64	10638.36
6/9/2018	17985.19	2115.48	4512.69	4958.62	10589.82
6/10/2018	17865.34	2096.07	4461.05	4894.55	10446.1
6/13/2018	17732.48	2079.06	4422.76	4848.44	10349.18
6/14/2018	17674.82	2075.32	4424.91	4843.55	10294.59
6/15/2018	17640.17	2071.5	4409.51	4834.93	10297.16
6/16/2018	17733.1	2077.99	4424.23	4844.91	10322.84
6/17/2018	17675.16	2071.22	4374.38	4800.34	10347.94
6/20/2018	17804.87	2083.25	4400.16	4837.21	10450.03
6/21/2018	17829.73	2088.9	4413.4	4843.76	10490.78
6/22/2018	17780.83	2085.45	4404.73	4833.32	10473.06
6/23/2018	18011.07	2113.32	4467.48	4910.04	10641.17
6/24/2018	17400.75	2037.41	4285.7	4707.98	10183.51
6/27/2018	17140.24	2000.54	4201.06	4594.44	9973.54
6/28/2018	17409.72	2036.09	4290.47	4691.87	10161.16
6/29/2018	17694.68	2070.77	4364.74	4779.24	10350.53
6/30/2018	17929.99	2098.86	4417.7	4842.67	10489.76
7/1/2018	17949.37	2102.95	4437.44	4862.57	10515.76
7/5/2018	17840.62	2088.55	4410.75	4822.9	10409.54
7/6/2018	17918.62	2099.73	4444.7	4859.16	10441.58
7/7/2018	17895.88	2097.9	4459.58	4876.81	10413.97
7/8/2018	18146.74	2129.9	4528.36	4956.76	10571.78
7/11/2018	18226.93	2137.16	4554.71	4988.64	10627.2
7/12/2018	18347.67	2152.14	4577.61	5022.82	10726.79
7/13/2018	18372.12	2152.43	4565.77	5005.73	10734.2
7/14/2018	18506.41	2163.75	4596.49	5034.06	10786.63
7/15/2018	18516.55	2161.74	4589.83	5029.59	10773.12
7/18/2018	18533.05	2166.89	4619.68	5055.78	10793.17
7/19/2018	18559.01	2163.78	4603.28	5036.37	10758.54
7/20/2018	18595.03	2173.02	4657.2	5089.93	10793.1
7/21/2018	18517.23	2165.17	4647.02	5073.9	10758.62
7/22/2018	18570.85	2175.03	4666.07	5100.16	10805.04
7/25/2018	18493.06	2168.48	4665.93	5097.63	10752.43
7/26/2018	18473.75	2169.18	4672.11	5110.05	10772.99

Продовження таблиці Б.1

7/27/2018	18472.17	2166.58	4702.89	5139.81	10739.77
7/28/2018	18456.35	2170.06	4721.41	5154.98	10744.16
7/29/2018	18432.24	2173.6	4730.23	5162.13	10785.51
8/1/2018	18404.51	2170.84	4756.04	5184.2	10730.2
8/2/2018	18313.77	2157.03	4719.21	5137.73	10661.13
8/3/2018	18355	2163.79	4734.33	5159.74	10695.14
8/4/2018	18352.05	2164.25	4743.81	5166.25	10707.13
8/5/2018	18543.53	2182.87	4791.21	5221.12	10782.87
8/8/2018	18529.29	2180.89	4784.74	5213.14	10788.01
8/9/2018	18533.05	2181.74	4795.75	5225.48	10804.51
8/10/2018	18495.66	2175.49	4783.39	5204.59	10774.98
8/11/2018	18613.52	2185.79	4803.31	5228.4	10836.15
8/12/2018	18576.47	2184.05	4807.08	5232.9	10822.41
8/15/2018	18636.05	2190.15	4827.12	5262.02	10858.5
8/16/2018	18552.02	2178.15	4797.19	5227.11	10809.74
8/17/2018	18573.94	2182.22	4805.45	5228.66	10824.64
8/18/2018	18597.7	2187.02	4808.68	5240.15	10863.01
8/19/2018	18552.57	2183.87	4806.14	5238.38	10829.15
8/22/2018	18529.42	2182.64	4808.56	5244.6	10815.92
8/23/2018	18547.3	2186.9	4818.48	5260.08	10847.49
8/24/2018	18481.48	2175.44	4783.48	5217.69	10791.18
8/25/2018	18448.41	2172.47	4775.43	5212.2	10784.37
8/26/2018	18395.4	2169.04	4783.69	5218.92	10749.33
8/29/2018	18502.99	2180.38	4791.09	5232.33	10811.35
8/30/2018	18454.3	2176.12	4775.99	5222.99	10797.54
8/31/2018	18400.88	2170.95	4771.06	5213.22	10764.75
9/1/2018	18419.3	2170.86	4783.94	5227.21	10771.93
9/2/2018	18491.96	2179.98	4798.74	5249.9	10856.92
9/6/2018	18538.12	2186.48	4829.51	5275.91	10890.79
9/7/2018	18526.14	2186.16	4831.86	5283.93	10890.18
9/8/2018	18479.91	2181.3	4804	5259.48	10882.33
9/9/2018	18085.45	2127.81	4681.53	5125.91	10613.53
9/12/2018	18325.07	2159.04	4764.65	5211.89	10741.24
9/13/2018	18066.75	2127.02	4722.85	5155.26	10535.36
9/14/2018	18034.77	2125.77	4746.06	5173.77	10511.4
9/15/2018	18212.48	2147.26	4819.9	5249.69	10602.94
9/16/2018	18123.8	2139.16	4818.05	5244.57	10532.27
9/19/2018	18120.17	2139.12	4796.07	5235.03	10564.3
9/20/2018	18129.96	2139.76	4805.22	5241.35	10560.26
9/21/2018	18293.7	2163.12	4853.75	5295.18	10708.37
9/22/2018	18392.46	2177.18	4891.36	5339.52	10793.66
9/23/2018	18261.45	2164.69	4858.91	5305.75	10717.99
9/26/2018	18094.83	2146.1	4817.17	5257.49	10624.88
9/27/2018	18228.3	2159.93	4866.71	5305.71	10665.34

Продовження таблиці Б.1

9/28/2018	18339.24	2171.37	4875.27	5318.55	10753.45
9/29/2018	18143.45	2151.13	4838.3	5269.15	10643.52
9/30/2018	18308.15	2168.27	4875.7	5312	10721.74
10/3/2018	18253.85	2161.2	4866.62	5300.87	10690.77
10/4/2018	18168.45	2150.49	4859.47	5289.66	10629.15
10/5/2018	18281.03	2159.73	4877.75	5316.02	10683.94
10/6/2018	18268.5	2160.77	4873.92	5306.85	10675.74
10/7/2018	18240.49	2153.74	4864.52	5292.4	10626.92
10/10/2018	18329.04	2163.66	4893.77	5328.67	10682.67
10/11/2018	18128.66	2136.73	4821.91	5246.79	10548
10/12/2018	18144.2	2139.18	4819.64	5239.02	10560.27
10/13/2018	18098.94	2132.55	4803.1	5213.33	10522.94
10/14/2018	18138.38	2132.98	4808.48	5214.16	10521.3
10/17/2018	18086.4	2126.5	4796.17	5199.82	10492.86
10/18/2018	18161.94	2139.6	4839.72	5243.84	10567.81
10/19/2018	18202.62	2144.29	4836.61	5246.41	10600.92
10/20/2018	18162.35	2141.34	4832.71	5241.83	10589.24
10/21/2018	18145.71	2141.16	4851.86	5257.4	10571.88
10/24/2018	18223.03	2151.33	4909.97	5309.83	10591.3
10/25/2018	18169.27	2143.16	4891.45	5283.4	10550.19
10/26/2018	18199.33	2139.43	4860.59	5250.27	10528.19
10/27/2018	18169.68	2133.04	4836.45	5215.97	10503.06
10/28/2018	18161.19	2126.41	4805.59	5190.1	10476.62
10/31/2018	18142.42	2126.15	4801.27	5189.14	10481.89
11/1/2018	18037.1	2111.72	4766.94	5153.58	10414.05
11/2/2018	17959.64	2097.94	4726.91	5105.57	10334.5
11/3/2018	17930.67	2088.66	4679.1	5058.41	10307.64
11/4/2018	17888.28	2085.18	4660.46	5046.37	10289.35
11/7/2018	18259.6	2131.52	4773.73	5166.17	10500.16
11/8/2018	18332.74	2139.56	4804.92	5193.49	10530.56
11/9/2018	18589.69	2163.26	4825.18	5251.07	10643.41
11/10/2018	18807.88	2167.48	4747.03	5208.8	10683.41
11/11/2018	18847.66	2164.45	4751.95	5237.11	10652.24
11/14/2018	18868.69	2164.2	4702.05	5218.4	10679.77
11/15/2018	18923.06	2180.39	4764.47	5275.62	10745.51
11/16/2018	18868.14	2176.94	4792.27	5294.58	10699.43
11/17/2018	18903.82	2187.12	4826.59	5333.97	10740.08
11/18/2018	18867.93	2181.9	4808.04	5321.51	10709.51
11/21/2018	18956.69	2198.18	4859	5368.86	10791.84
11/22/2018	19023.87	2202.94	4873.84	5386.35	10820.18
11/23/2018	19083.18	2204.72	4853.86	5380.68	10835.91
11/25/2018	19152.14	2213.35	4870.02	5398.92	10878.09
11/28/2018	19097.9	2201.72	4857.02	5368.81	10808.63
11/29/2018	19121.6	2204.66	4873.21	5379.92	10817.7

Продовження таблиці Б.1

11/30/2018	19123.58	2198.81	4810.81	5323.68	10838.46
12/1/2018	19191.93	2191.08	4734.1	5251.11	10828.99
12/2/2018	19170.42	2191.95	4739.37	5255.65	10838.58
12/5/2018	19216.24	2204.71	4778.14	5308.89	10910.9
12/6/2018	19251.78	2212.23	4788.67	5333	10970.78
12/7/2018	19549.62	2241.35	4851.94	5393.76	11114.61
12/8/2018	19614.81	2246.19	4859.19	5417.36	11149.96
12/9/2018	19756.85	2259.53	4895.9	5444.5	11191.79
12/12/2018	19796.43	2256.96	4874.3	5412.54	11177.28
12/13/2018	19911.21	2271.72	4935.84	5463.83	11237.17
12/14/2018	19792.53	2253.28	4921.22	5436.67	11098.67
12/15/2018	19852.24	2262.03	4933.41	5456.85	11131.89
12/16/2018	19843.41	2258.07	4914.86	5437.16	11125.22
12/19/2018	19883.06	2262.53	4934.85	5457.44	11128.54
12/20/2018	19974.62	2270.76	4953.8	5483.94	11172.19
12/21/2018	19941.96	2265.18	4948.9	5471.43	11142.57
12/22/2018	19918.88	2260.96	4934.39	5447.42	11114.14
12/23/2018	19933.81	2263.79	4940.02	5462.69	11128.8
12/27/2018	19945.04	2268.88	4965.81	5487.44	11146.4
12/28/2018	19833.68	2249.92	4926.29	5438.56	11058.88
12/29/2018	19819.78	2249.26	4918.28	5432.09	11074.33
12/30/2018	19762.6	2238.83	4863.62	5383.12	11056.89
1/3/2019	19881.76	2257.83	4911.33	5429.08	11154.35
1/4/2019	19942.16	2270.75	4937.2	5477.01	11246.54
1/5/2019	19899.29	2269	4964.95	5487.94	11247.69
1/6/2019	19963.8	2276.98	5007.08	5521.06	11237.62
1/9/2019	19887.38	2268.9	5024.9	5531.82	11169.79
1/10/2019	19855.53	2268.9	5035.17	5551.82	11183.33
1/11/2019	19954.28	2275.32	5050.21	5563.65	11226.78
1/12/2019	19891	2270.44	5041.43	5547.49	11204.15
1/13/2019	19885.73	2274.64	5059.51	5574.12	11227.17
1/17/2019	19826.77	2267.89	5044.65	5538.73	11196.29
1/18/2019	19804.72	2271.89	5055.85	5555.65	11196.1
1/19/2019	19732.4	2263.69	5051.17	5540.08	11148.85
1/20/2019	19827.25	2271.31	5063.2	5555.33	11192.79
1/23/2019	19799.85	2265.2	5065.7	5552.94	11170.62
1/24/2019	19912.71	2280.07	5101.06	5600.96	11249.28
1/25/2019	20068.51	2298.37	5151.47	5656.34	11339.05
1/26/2019	20100.91	2296.68	5156.92	5655.18	11313.12
1/27/2019	20093.78	2294.69	5168.06	5660.78	11283.19
1/30/2019	19971.13	2280.9	5129.33	5613.71	11205.24
1/31/2019	19864.09	2278.87	5116.77	5614.79	11222.95
2/1/2019	19890.94	2279.55	5152.69	5642.65	11207.24
2/2/2019	19884.91	2280.85	5147.7	5636.2	11215.38

Продовження таблиці Б.1

2/3/2019	20071.46	2297.42	5161.6	5666.77	11310.74
2/6/2019	20052.42	2292.56	5167.96	5663.55	11264.11
2/7/2019	20090.29	2293.08	5185.88	5674.22	11243.38
2/8/2019	20054.34	2294.67	5196.58	5682.45	11251.8
2/9/2019	20172.4	2307.87	5212.16	5715.18	11327.68
2/10/2019	20269.37	2316.1	5226.69	5734.13	11377.72
2/13/2019	20412.16	2328.25	5256.82	5763.96	11433.93
2/14/2019	20504.41	2337.58	5271.07	5782.57	11468.87
2/15/2019	20611.86	2349.25	5302.39	5819.44	11514.4
2/16/2019	20619.77	2347.22	5300.64	5814.9	11504.19
2/17/2019	20624.05	2351.16	5324.72	5838.58	11510.91
2/21/2019	20743	2365.38	5350.73	5865.95	11578.25
2/22/2019	20775.6	2362.82	5352.13	5860.63	11550.18
2/23/2019	20810.32	2363.81	5332.37	5835.51	11556.15
2/24/2019	20821.76	2367.34	5343.31	5845.31	11541.29
2/27/2019	20837.44	2369.75	5347.55	5861.9	11558.35
2/28/2019	20812.24	2363.64	5330.31	5825.44	11512.39
3/1/2019	21115.55	2395.96	5390.99	5904.03	11661.22
3/2/2019	21002.97	2381.92	5363.26	5861.22	11575.91
3/3/2019	21005.71	2383.12	5373.48	5870.75	11598.37
3/6/2019	20954.34	2375.31	5360.31	5849.18	11547.79
3/7/2019	20924.76	2368.39	5351.28	5833.93	11506.32
3/8/2019	20855.73	2362.98	5359.81	5837.55	11448.21
3/9/2019	20858.19	2364.87	5363.98	5838.81	11457.64
3/10/2019	20902.98	2372.6	5385.9	5861.73	11500.76
3/13/2019	20881.48	2373.47	5394.57	5875.78	11519.44
3/14/2019	20837.37	2365.45	5382.17	5856.82	11459.29
3/15/2019	20950.1	2385.26	5416.25	5900.05	11590.31
3/16/2019	20934.55	2381.38	5412.08	5900.76	11600.23
3/17/2019	20914.62	2378.25	5408.76	5901	11589.01
3/20/2019	20905.86	2373.47	5413.31	5901.53	11556.93
3/21/2019	20668.01	2344.02	5332.53	5793.82	11420.43
3/22/2019	20661.3	2348.45	5367.56	5821.64	11426.69
3/23/2019	20656.58	2345.96	5355.14	5817.69	11430.89
3/24/2019	20596.72	2343.98	5363.99	5828.74	11418.89
3/27/2019	20550.98	2341.59	5374.27	5840.37	11414.33
3/28/2019	20701.5	2358.57	5407.21	5875.14	11493.84
3/29/2019	20659.32	2361.13	5430.27	5897.55	11496.67
3/30/2019	20728.49	2368.06	5439.74	5914.34	11519.84
3/31/2019	20663.22	2362.72	5436.23	5911.74	11492.85
4/3/2019	20650.21	2358.84	5432.2	5894.68	11463.91
4/4/2019	20689.24	2360.16	5440.41	5898.61	11470.54
4/5/2019	20648.15	2352.95	5418.52	5864.48	11423.36
4/6/2019	20662.95	2357.49	5420.88	5878.95	11457.29

Продовження таблиці Б.1

4/7/2019	20656.1	2355.54	5418.37	5877.81	11445.58
4/10/2019	20658.02	2357.16	5421.68	5880.93	11464.34
4/11/2019	20651.3	2353.78	5398.4	5866.77	11473.62
4/12/2019	20591.86	2344.93	5377.03	5836.16	11423.17
4/13/2019	20453.25	2328.95	5353.59	5805.15	11324.53
4/17/2019	20636.92	2349.01	5399.2	5856.79	11427.08
4/18/2019	20523.28	2342.19	5391.66	5849.47	11378.58
4/19/2019	20512.19	2349.35	5399.64	5863.03	11342.42

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Рівень вищої освіти – другий (магістерський)

Спеціальність (спеціалізація) – 122 «Комп'ютерні науки» («Інтелектуальний аналіз даних в управлінні проектами»)

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри ММСА

Тимощук О.Л.

«___»_____2019 р.

ЗАВДАННЯ
на магістерську дисертацію студенту
Пінчуку Владиславу Олександровичу

1. Тема дисертації «Моделювання нелінійних нестационарних процесів методами інтелектуального аналізу даних», науковий керівник дисертації Бідюк Петро Іванович, д.т.н., проф., професор кафедри ММСА ННК «ІПСА», затверджені наказом по університету від «___»_____ 20__ р. №_____

2. Термін подання студентом дисертації _____

3. Об'єкт дослідження: нестационарні процеси різної природи в економіці та фінансах.

4. Предмет дослідження: інтелектуальні методи моделювання та прогнозування нелінійних нестационарних процесів.

5. Перелік завдань, які потрібно розробити:

- (1) дослідити актуальність задач моделювання та прогнозування сьогодні;
- (2) виконати огляд сучасних моделей нестационарних процесів;
- (3) обрати та описати методи прогнозування на основі інтелектуального аналізу даних;
- (4) виконати обчислювальні експерименти стосовно моделювання вибраних процесів та обчислення оцінок прогнозів;

- (5) виконати аналіз та порівняння отриманих результатів;
- (6) дослідити можливі шляхи покращення результатів;
- (7) дослідити доцільність побудови СППР для розв'язання задач моделювання і прогнозування фінансово-економічних процесів;
- (8) розробити план для створення стартап-проекту.
6. Орієнтовний перелік графічного (ілюстративного) матеріалу: (1) – постановка задачі дослідження; (2) – огляд використаних методів моделювання і прогнозування; (3) – критерії для аналізу адекватності моделей і якості оцінок прогнозів; (4) – результати моделювання і прогнозування вибраних процесів.
7. Орієнтовний перелік публікацій:
8. Дата видачі завдання _____

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Пошук літератури та поверхневе дослідження за темою	12.03.2018 – 30.04.2018	Виконано
2	Огляд сучасних методів інтелектуального аналізу даних що застосовуються для моделювання і прогнозування процесів	01.05.2018 – 30.05.2018	Виконано
3	Створення системи підтримки прийняття рішень для моделювання і прогнозування	01.06.2018 – 30.07.2018	Виконано
4	Опис СППР	01.09.2018 – 20.09.2018	Виконано
5	Пошук та попередня підготовка даних для виконання обчислювальних експериментів	22.09.2018 – 10.10.2018	Виконано
6	Виконання обчислювальних експериментів з метою побудови моделей і короткострокового прогнозування. Аналіз адекватності побудованих моделей і отриманих оцінок прогнозів.	11.10.2018 – 14.12.2018	Виконано
7	Аналіз можливостей подальшого покращення оцінок прогнозів.	15.12.2018 – 30.12.2018	Виконано
8	Розробка плану запуску стартап-проекту	10.01.2019 – 28.02.2019	Виконано
9	Підготовка ілюстративних матеріалів для захисту роботи.	01.03.2019 – 30.03.2019	Виконано
10	Формування звіту дисертації та її захист.	01.04.2019 – 18.05.2019	Виконано

Студент

В.О. Пінчук

Науковий керівник дисертації

П.І. Бідюк